Structural and Functional Analysis of the Levansucrase SacB from *Bacillus megaterium*

Von der Fakultät für Lebenswissenschaften

der Technischen Universität Carolo-Wilhelmina

zu Braunschweig

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

genehmigte

D i s s e r t a t i o n

von Christian Philipp Strube
aus Tübingen

1. Referent: Honorarprofessor Dr. Dirk Heinz
2. Referentin: Dr. Gunhild Layer
eingereicht am: 07.03.2011
mündliche Prüfung (Disputation) am 01.07.2011

Druckjahr 2011

Vorveröffentlichungen der Dissertation

Teilergebnisse aus dieser Arbeit wurden mit Genehmigung der Fakultät für Lebenswissenschaften, vertreten durch den Mentor der Arbeit, in folgenden Beiträgen vorab veröffentlicht:

Publikationen

Strube, C. P., Homann, A., Gamer, M., Jahn, D., Seibel, J. Heinz, D. W.: Polysaccharide synthesis of the Levansucrase SacB from *Bacillus megaterium* is controlled by distinct surface motifs.

Tagungsbeiträge

Strube, C. P., Timm, M., Homann, A., Memmel, E., Seibel, J. & Heinz, D. W.: Strukturbiologie von Glycosyltransferasen zur Optimierung von biotechnologischen Prozessen. (Poster), 100. Kolloquium des Sonderforschungsbereiches 578, Braunschweig (2010).

Strube, C. P., Homann, A. Götze, S., Biedendieck, R., Seibel, J. & Heinz, D. W.: X-ray Analysis and Biochemical Characterization of SacB, a Novel Fructosyltransferase from *Bacillus megaterium*. (Poster), International Congress on Biocatalysis, Hamburg (2008).

Strube, C. P., Homann, A., Biedendieck, R., Gamer, M., Jahn, D., Heinz, D. W. & Seibel, J.: Strukturbiologie von Glykosyltransferasen zur Optimierung von biotechnologischen Prozessen. (Poster), Begutachtung SFB 578, Braunschweig (2008).

Strube, C. P., Homann, A., Biedendieck, R., Jahn, D., Heinz, D. W. & Seibel, J.: Strukturbiologie von Glykosyltransferasen zur Optimierung von biotechnologischen Prozessen. (Vortrag), Berichtskolloquium SFB 578, Braunschweig (2007).

Strube, C. P., Homann, A. Götze, S., Biedendieck, R., Seibel, J. & Heinz, D. W.: Purification, crystallization and preliminary X-ray analysis of SacB, a fructosyltransferase from *Bacillus megaterium*. (Poster), European BioPerspectives, Köln (2007).

Homann, A., Strube, C. P., Pasche, B., Schughart, K., Heinz, D. W. & Seibel, J.: Enzymatic synthesis and immuno-assay of novel carbohydrate structures. (Poster), EMBL International PhD Symposium, Heidelberg (2009).

Homann, A., Strube, C. P., Pasche, B., Schughart, K., Heinz, D. W. & Seibel, J.: Towards tailor-made oligosaccharides – Chemo-enzymatic synthesis and physiological functions of novel carbohydrate structures. (Poster, Oral Presentation), SFB 630 International Symposium, Novel agents against Infectious diseases – An Interdisciplinary approach, Würzburg (2009).

Homann, A., Strube, C. P., Jahn, D., Heinz, D. W. & Seibel, J.: Towards tailor-made oligosaccharides by enzyme and substrate engineering. (Poster Award), 15th European Carbohydrate Symposium, Wien (2009).

Homann, A., Zuccaro, A. Götze, S., Strube, C. P., Dersch, P., Heinz, D. W. & Seibel, J.: Novel fructo-oligosaccharides as pharma-/nutraceuticals. (Poster), European Bioperspectives, Hannover (2008).

Homann, A., Strube, C. P., Heinz, D. W. & Seibel, J.: Oligo- *vs*. polysaccharide formation – Production of oligofructosides and structure-function analysis of a novel fructosyltransferase from *Bacillus megaterium.* (Poster), International Congress on Biocatalysis, Hamburg (2008).

CONTENTS

Abbreviations

Å	Ångstrøm (1 Å = 0.1 nm)
A_λ	Absorption at the wavelength λ in nm (equivalent to OD_λ)
Amp	Ampicillin
Amp^R	Ampicillin resistance
BESSY	Berliner Elektronenspeicherring-Gesellschaft für Synchrotron Strahlung, Berlin, Germany
BLAST	Basic Local Alignment Search Tool
B. megaterium	*Bacillus megaterium*
C-	Carboxy terminus
Ccp4	Collaborative Computational Project 4
CHT	Ceramic hydroxyapatite
Cm	Chloramphenicol
Cm^R	Chloramphenicol resistance
CV	Column volume
Da	Dalton (equals the mass of 1/12 of the carbon ^{12}C isotope)
DALI	Distance Alignment Server
DESY	Deutsches Elektronensynchrotron, Hamburg, Germany
DLS	Dynamic Light Scattering
DNA	Deoxyribonucleic acid
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
EMBL	European molecular biology laboratory
EPS	Exopolysaccharides
ESRF	European synchrotron radiation facility, Grenoble, France
EtOH	Ethanol
GBD	Glucan binding domain
GLRF	General Locked Rotation Function
GtfR	Glycosyltransferase R
His_6	Six successive histidine residues, used as affinity tag
HPAEC	High Performance Anion Exchange Chromatography
IEC	Ion exchange chromatography
IPTG	Isopropyl *β*-D-thiogalactopyranoside
Kan	Kanamycin
Kan^R	Kanamycin resistance
kDa	Kilodalton
LAB	lactic acid bacteria
LB	Lysogeny broth
MeOH	Methanol
M_r	Molecular mass
MR	Molecular Replacement
MS	Mass spectrometry
MW	Molecular weight

MWCO	Molecular weight cut-off
N-	Amino terminus
Ni-NTA	Nickel (II) nitrilotriacetic acid
o.n.	Overnight
OD_λ	Optical density at the wavelength λ in nm
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank
PEG	Polyethylene glycol
PVDF	Polyvinylidene difluoride
RADAR	Rapid Automatic Detection and Alignment of Repeats
r.m.s.d.	Root mean square deviation
rpm	Rotations per minute
RT	Room temperature
σ	Standard deviation
S. oralis	*Streptococcus oralis*
SDS	Sodium dodecyl sulfate
TEMED	N,N,N',N'-Tetramethylethylenediamine
TEV	Tobacco etch virus
TIM	Triose phosphate isomerase
TIM barrel	Protein fold first described in triose phosphate isomerase
TLC	Thin layer chromatography
Tris	Tris-(hydroxymethyl) aminomethane
TU	Technical University
V_M	Matthews coefficient
wt	Wildtype

ZUSAMMENFASSUNG

Poly- und Oligosaccharide bakteriellen Ursprungs finden sowohl in der Biomedizin als auch in der Nahrungsmittelindustrie vielfach ihre Anwendung. Die Synthese dieser komplexen Kohlenhydratstrukturen stellt jedoch eine Herausforderung für die moderne Wissenschaft dar. Katalysiert wird ihre Biosynthese durch Enzyme, die als Glykosyltransferasen bezeichnet werden. Trotz der großen biologischen Bedeutung und des enormen pharmazeutischen Potentials der Saccharide ist der katalytische Mechanismus der Enzyme zur Bildung von Poly- und Oligosacchariden weitestgehend unklar.

SacB aus *Bacillus megaterium*

Die Levansucrase SacB aus *B. megaterium* synthetisiert aus Saccharose hochmolekulares β-(2,6) verknüpftes Levan durch die Übertragung von Fructosyl-Resten. In der vorliegenden Arbeit wurden die Kristallstrukturen der SacB Varianten Y247A, Y247W, N252A, D257A und K373A bei Auflösungen zwischen 1.75 und 2.0 Å bestimmt. Anhand von SacB konnte zum ersten Mal gezeigt werden, dass Aminosäuren, die sich außerhalb des aktiven Zentrums befinden, einen definierten Einfluss auf die Synthese von Polysacchariden bei Glycosyltransferasen haben. Die ermittelten Daten der Strukturanalyse korrelieren zudem mit den Ergebnissen der kinetischen und biochemischen Untersuchungen. Die Architektur des aktiven Zentrums der verschiedenen Varianten ist unverändert und intakt. Diese Beobachtung wird durch die kinetischen Parameter der Varianten bestätigt, die nur leicht von den Werten des Wildtypenzyms abweichen. Darüber hinaus deuten die Kristallstrukturen auf ein mögliches Arrangement von Aminosäuren für die Bindung einer Akzeptor-Fructosylkette auf der Oberfläche des Enzyms hin. Die Aminosäuren Asn252, Lys373 und Tyr247 bilden eine mögliche Plattform für die Stabilisierung der wachsenden Fructosylkette. Eindeutige Bindungsstellen für die Zuckerkette konnten den ausgetauschten Aminosäuren zugeordnet werden. Die Varianten K373A, N252A und Y247A synthetisieren jeweils

definierte Mischungen aus Oligosacchariden. Die klar unterscheidbaren Kettenlängen korrelieren mit den zugeordneten Bindungsstellen auf der Oberfläche des Enzyms.
Diese Ergebnisse tragen zu einem besseren Verständnis des Mechanismus zur Übertragung von Kohlenhydraten bei. Trotz dieses Beitrags mangelt es jedoch weiterhin an Komplexstrukturen aus Enzym und gebundenen Oligosacchariden, die eine noch präzisere Beschreibung dieser Prozesse erlauben würden.

GtfR aus *Streptococcus oralis*

Die Glycosyltransferase R (GtfR) aus *S. oralis* ist eine Dextran produzierende Glucansucrase. Verschiedene Studien haben gezeigt, dass die C-terminale Domäne dieser Enzyme an der Bindung von Glucan beteiligt ist und darüber hinaus für die Aktivität des Enzyms benötigt wird. In der vorliegenden Arbeit wurde die glucanbindende Domäne (GBD) der GtfR erfolgreich in *E. coli* produziert und anschließend gereinigt. Versuche, die GBD zu kristallisieren, ergaben kleine Kristalle, die für eine Röntgenanalyse nicht geeignet waren. Eine Optimierung dieser ersten Kristalle mittels *microseeding* resultierte in deutlich größeren Kristallen, die in Röntgenexperimenten jedoch ein stark anisotropes Diffraktionsmuster aufwiesen. In der einen Richtung lag die Auflösung bei Werten zwischen 3.6 und 3.9 Å, in der anderen Richtung zwischen 5.8 und 6.2 Å. Die Bemühungen, die Struktur über die Methode des molekularen Ersatz zu lösen, waren nicht erfolgreich. Aufgrund der anisotropen Daten und der nur geringen Auflösung wurde die Strukturbestimmung nicht weiter verfolgt. Für erfolgreiche Arbeiten wird in Zukunft eine signifikante Verbesserung der Kristallqualität notwendig sein. Die hier erzielten Ergebnisse können jedoch als Ausgangspunkt für weitere Optimierungen der Kristalle genutzt werden.

SUMMARY

Bacterial poly- and oligosaccharides find their applications in various fields ranging from biomedical applications to the use in the food and non-food industry. However, the synthesis of poly- and oligosaccharides still represents a challenging task for the modern life sciences. The biosynthesis is catalyzed by the action of enzymes commonly referred to as glycosyltransferases. Despite the widespread biological importance of such carbohydrates and their enormous pharmaceutical potential, the polysaccharide synthesis mechanisms of glycosyltransferases are currently underexplored.

SacB from *Bacillus megaterium*

Levansucrase SacB from *B. megaterium* synthesizes high molecular weight, β-(2,6) linked levan from sucrose by transfer of fructosyl units. In this study, the structures of SacB variants Y247A, Y247W, N252A, D257A and K373A at resolutions between 1.75 and 2.0 Å have been determined. By means of SacB, it is shown for the first time that amino acids remote to the active site of a polysaccharide-forming enzyme have a well-defined and rationally explainable effect on the sugar polymer formation activity. The structural data is consistent with the kinetic and biochemical analyses. Conformational analyses of variants Y247A, Y247W, N252A and K373A reveal an unchanged active site architecture. Supporting the crystallographic data, the kinetic parameters of these variants are not significantly different compared to the wild-type SacB. Moreover, the structural data point towards a possible surface arrangement for the binding of an acceptor fructosyl chain. Residues Asn252, Lys373 and Tyr247 form a platform for a possible stabilization of the acceptor fructan chain. Defined subsites can be assigned to certain amino acids. Variants K373A, N252A and Y247A synthesize unique mixtures of oligosaccharides of clearly distinguishable and different chain lengths correlating to their subsites on the surface of SacB.

These results lead to a better understanding of carbohydrate transfer mechanisms. However, structures of complexes between SacB and oligosaccharides of different length are still missing. They would provide a more comprehensive explanation of the

transfructosylation process regarding oligo- and polysaccharide synthesis.

GtfR from *Streptococcus oralis*

Glycosyltransferase R (GtfR) from *S. oralis* is a dextran-producing glucansucrase. Several studies demonstrated that the C-terminal domain of these enzymes is involved in glucan binding and is necessary to keep an active enzyme. In this thesis, the glucan binding domain (GBD) of the GtfR has been successfully produced in *E. coli* and purified. Attempts to crystallize the GBD yielded in small crystals not suitable for X-ray analysis. Optimization of the initial crystals by microseeding resulted in larger crystals. However, these crystals gave only highly anisotropic diffraction patterns with spatial resolutions of 3.6 – 3.9 Å in one direction and 5.8 – 6.2 Å in the other direction. Efforts to determine the structure by molecular replacement were not successful. Due to the anisotropic and low resolution data, structure solution of the GBD was finally abandoned. For future work, generation of high quality diffracting crystals will be mandatory. The reported results can however be regarded as a starting point for further optimization of GBD crystals.

1 INTRODUCTION

1.1 Carbohydrates

For a long time the central dogma of modern biology was that biological information flows from DNA to RNA to protein. The quality of this concept lies in the template-based precision and in particular in the ability to manipulate any one class of molecules based on knowledge of the other (Varki *et al*., 1999). The patterns of sequence homology and relatedness allowed predicting function and revealing evolutionary relationships. However, as our knowledge of transcriptional and translation control and regulation improved as well as the successive discovery of complex co- and posttranslational modifications such as glycosylation of proteins, it became apparent that an extension of this paradigm is necessary. Creating a cell requires, besides DNA, RNA and proteins two further major classes of molecules: carbohydrates and lipids.

Carbohydrates ($C_n(H_2O)_n$) represent one of the three major classes of biological macromolecules along with proteins and nucleic acids and make up most of the organic matter on earth (Berg *et al*., 2002). The simplest carbohydrates are called monosaccharides, which cannot be hydrolyzed into simpler units. They can be distinguished into two different types: aldoses and ketoses. Their predominant form is the ring form, leading to the formation of an anomeric center. Monosaccharides are the basic structural unit for all higher carbohydrates. Glycosidic linkages involving the free hydroxyl group of the anomeric center leads to the formation of linear or branched chains of monosaccharides, called oligo- or polysaccharides, which show remarkable structural variations. Sucrose, for example is a non-reducing heterodisaccharide composed of α-D-glucopyranose and β-D-fructofuranose (Figure 1-1).

Carbohydrates can be divided into structural and non-structural components of cells (van Hijum, 2006). Structural carbohydrates are located within the cell-wall and the cytoskeleton. Examples are cellulose, chitin, murein, and pectins. Non-structural carbohydrates are mostly known as energy-rich compounds of the metabolism and as energy storage molecules. Fructose and glucose are non-structural carbohydrates found

naturally in fruits and vegetable. They represent major products of the photosynthesis. The huge diversity of carbohydrates, comprising the degree of polymerisation, the linkage types and the anomeric state of carbohydrates are factors making the field of glycobiology so complex and challenging.

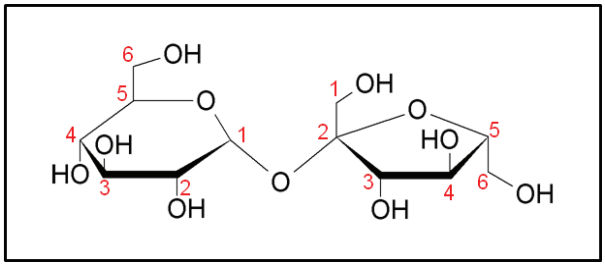

Figure 1-1: Sucrose molecule consisting of an α-D-glucopyranosyl unit linked 1→2 to an β-D-fructofuranoside. The C-atoms are numbered in red.

Exopolysaccharides

The cell surface of microbial organisms is a rich source of carbohydrate-containing molecules (Varki, 1999). In addition to microbial cell wall components such as bacterial teichoic acid, lipopolysaccharides and peptidoglycan, polysaccharides may be found associated with other surface macromolecules or totally dissociated from the microbial cell (Sutherland, 1990; Medzhitov and Janeway, 1997). These extracellular polysaccharides are named exopolysaccharides (EPS) and are produced by a great variety of bacteria. Most microalgae yield some type of EPS too. Among yeast and fungi they are less frequently found (Parolis *et al.*, 1998). Bacterial EPS have varying length and can reach a size between 10 to 10^4 kDa (approximately 50 to 50,000 glycosyl units). They show a considerable diversity in their composition and structure, with respect to the type of linkages (Sutherland, 1990). Depending on their monosaccharide composition and the mechanism of biosynthesis, bacterial EPS can be divided into two classes: heteropolysaccharides and homopolysaccharides (de Vuyst and Degeest, 1999; Jolly *et al.*, 2002).

Bacterial heteropolysaccharides

Heteropolysaccharides are polymers comprising two or more types of monosaccharides, mostly glucose, fructose, galactose and rhamnose (De Vuyst and Vaningelgem, 2003).

For the majority of known polysaccharides various organic or inorganic substituents such as sulfate, phosphate and acyl substituents are present (Kenne and Lindberg, 1983; Sutherland, 1990; Geresh *et al.*, 2002). The biosynthesis of heteropolysaccharides resembles the biosynthetic pathways of the bacterial cell wall components peptidoglycan, lipopolysaccharide and teichoic acid (Kumar *et al.*, 2007). The intracellular synthesis occurs within the cytoplasm by the combined action of several enzymes named glycosyltransferases. The synthesized exopolysaccharides are exported to final sites remote to the cytoplasmic membrane (Sutherland, 1990). Heteropolysaccharides are predominantly produced by lactic acid bacteria such as *Lactobacillus sake* (Van den Berg *et al*., 1993), *Bifidobacterium adolescentis* (Hosono *et al*., 1997), *Streptococcus salivarius ssp. thermophilus* (Stingele *et al*., 1996), *Lactobacillus delbrueckii ssp. bulgaricus* (Cerning *et al*., 1986; Grobben *et al*., 1997) and various *Lactococcus* strains (Cerning *et al*., 1990; Kojic *et al*., 1992; Van Casteren *et al*., 2000).

Bacterial homopolysaccharides

Bacterial homopolysaccharides are composed of just one type of monosaccharide. In general two different types of homopolysaccharides are known. Either they consist of fructose units or of glucose units. The respective polysaccharides are named fructans and glucans. They are synthesized in the form of slime on the outside of the microbial cells by enzymes called sucrases (Figure 1-2) (van Hijum, 2006).

Figure 1-2: ***Lactobacillus reuteri* 121 bacteria growing on agar supplemented with sucrose**. The produced exopolysaccharides form a slimy, sticky layer surrounding the colonies. (Vujičić-Žagar, 2007).

In contrast to the synthesis of heteropolysaccharides, which requires multiple enzymes and nucleotide-activated sugars, homopolysaccharides are usually synthesized by the action of just one enzyme.

Fructans

Fructans are soluble polymers of β-D-fructofuranose. Essentially, there are two kinds of fructans, levan and inulin. Fructans of the levan and the inulin type produced by bacteria are mostly of high molecular weight with molecular masses up to 50 x 10^6 kDa (Chambert and Petit-Glatron, 1991; van Hijum *et al*., 2006). Levans and inulins are linear polymers which differ in the glycosidic bond between their fructose units. Levan is a polysaccharide containing mainly β-(2,6) linked units of D-fructofuranose (Figure 1-3). Every one to four units β-(2,1) linked branches of D-fructofuranose onto the main chain occur (Robyt, 1998). The addition of a fructosyl residue to sucrose with a β-(2,6) bond results in the formation of 6-kestose, the first intermediate of levan biosynthesis (Velázquez-Hernández *et al*., 2008) The polymer inulin contains D-fructofuranosyl units linked β-(2,1) with branches of β-(2,6) linked units (Figure 1-3). Fructan synthesis by several bacillus strains has been reported, i.e. *B. salivarius*, *B. subtilis* and *B. amyloliquefaciens* (Chambert *et al*., 1974; Macura, and Townsley, 1984; Song and Jacques, 1999). For all of these bacteria the produced fructans have been identified as levan polymers (van Hijum *et al,* 2006).

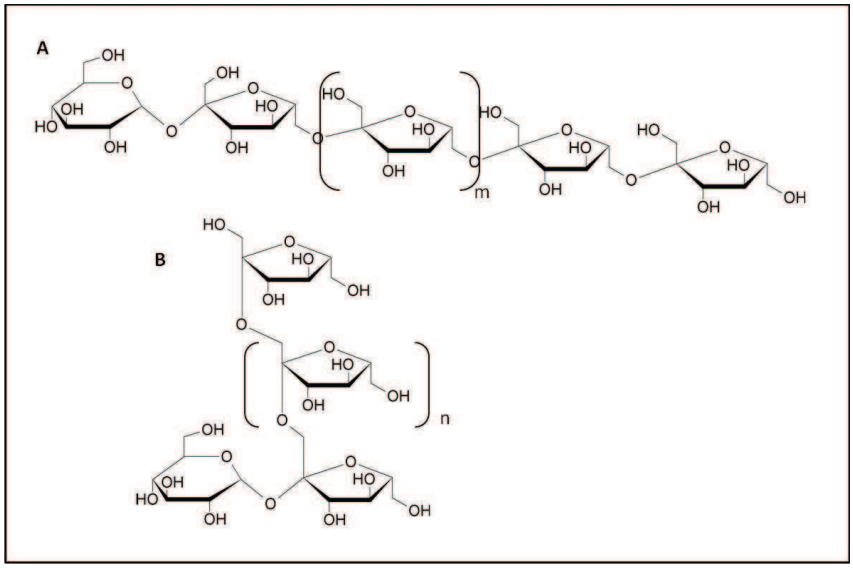

Figure 1-3: **Chemical structures of fructans. A**: section of levan (where n > 100); **B**: section of inulin (where m = 20-10000).

Glucans

Glucans are polymers of α-D-glucopyranosyl residues, synthesized by glucansucrases. Depending on the glucansucrase-producing strain, four different types of glucans with

different sizes and structure are synthesized (Monchois *et al*., 1999). A well-known example is the water-soluble polymer dextran with α-(1,6) glucosidic bonds in the main chain (Figure 1-4) (Cerning, 1990). Polyglucose mutan is a water-insoluble glucan and contains predominantly α-(1,3) glucopyranosyl units (Figure 1-4) (Hamada and Slade, 1980). The alternan polymer is linked alternately through α-(1,6) and α-(1,3) bonds (Arguello-Morales *et al*., 2000), while reuteran contains mainly α-(1,4) glucosidic bonds and significant amounts of α-(1,6) and α-(1,4,6) glucosidic linkages (van Geel-Schutten *et al*., 1999; Kralj *et al*., 2002).

Differences in length, the type of linkages and the type and degree of branches result in a large variation in solubility and other physical characteristics of the glucans.

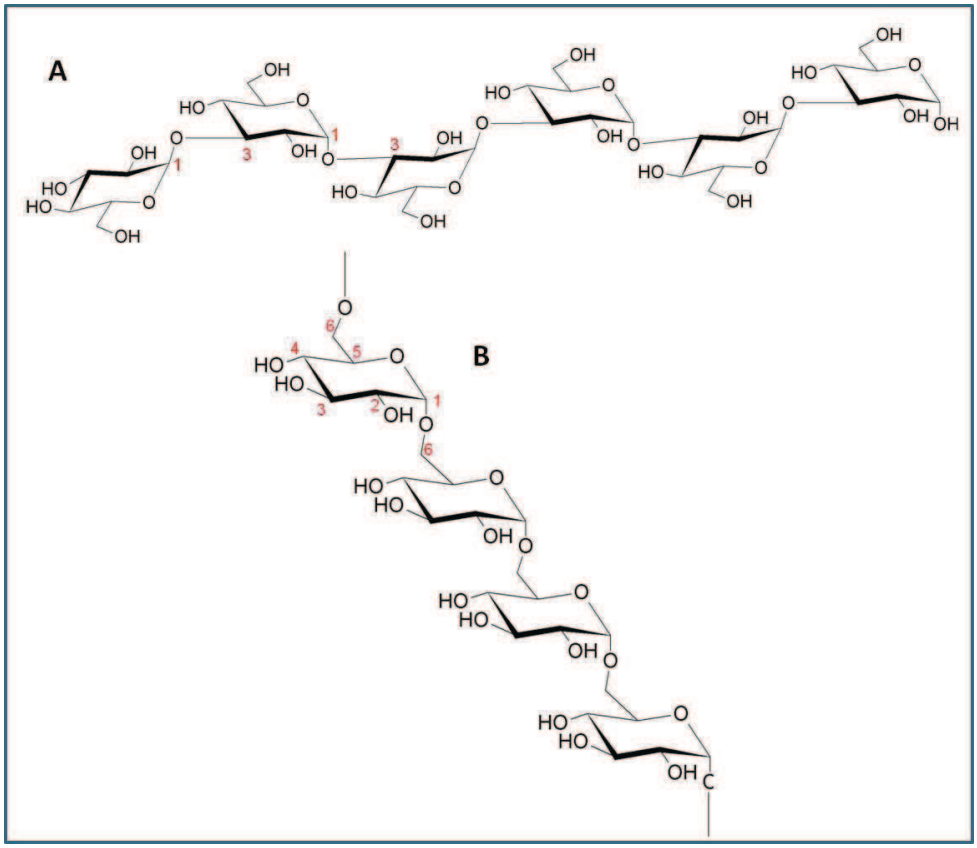

Figure 1-4: **Chemical structures of glucans**: A: section of mutan; B: section of dextran.

Oligosaccharides

Oligosaccharides contain only a small number of linked monosaccharides, mostly with a degree of polymerization ranging from 2 - 6 units. Under certain circumstances the same

enzymes synthesizing high-molecular-mass fructans and glucans produce such oligosaccharides instead of long-chain polymers. These fructo- and gluco-oligosaccharides may contain several types of linkages between the fructose and glucose units (van Hijum, 2004). Depending on the enzyme and the reaction conditions, a substantial number of different oligosaccharides is obtainable. They are often found as components of glycoproteins or glycolipids, which are involved in processes like inflammation, cell-cell signaling and infections (Ungar, 2009). A well-known example is the human blood system. The structures of the blood types A and B are highly similar with the exception of one *N*-linked acetyl group of the terminal carbohydrate unit. The blood type 0 is characterized by the absence of the terminal galactose or *N*-acetylgalactosamine, respectively. A small change in the structure of the tetrasaccharide has a dramatic effect and antibodies will attack the antigens on any other blood type.

1.2 Biological role of microbial exopolysaccharides

Only little is known about the biological role of EPS in bacteria. It is assumed that there are multiple functions which depend on the environment of the microorganisms. Most of the functions ascribed are of a protective nature. Microbial cells surrounded by a highly hydrated layer of EPS may be provided with protection against antibiotics or toxic compounds (e.g. metal ions, sulfur dioxide and ethanol) or against desiccation and osmotic stress due to the excellent water-binding properties of EPS. Capsular EPS can inhibit predation by protozoans, phagocytosis and phage attack (Sutherland, 1972; Kang and Cottrell, 1979; Cerning, 1990; Roberts, 1995; Jolly *et al*., 2002). Moreover, EPS have been recognized in adherent biofilms and biological surfaces, suggesting a role in the formation of dental plaque on tooth surfaces by *Streptococcus sp*. (Parker and Creamer, 1971; Whitfield and Keenleyside, 1995; Kralj, 2004).

It is further known that complex carbohydrates are also implicated in a number of cellular recognition reactions. Therefore EPS are assumed to be involved in host-pathogen interactions (Jolly and Stingele, 2001; Broadbent *et al*., 2003).

The role of EPS as extracellular energy reserves has also been reported, but the vast

majority of EPS forming bacteria is lacking degrading enzymes that is not being able to metabolize the produced polysaccharides (Cerning, 1990; van Hijum, 2004).

1.3 Applications of poly- and oligosaccharides

Bacterial poly- and oligosaccharides find their applications in various fields ranging from medicine to the use in the food as well as in the non-food industry (Sutherland, 1990). A variety of polysaccharides like hyaluronan, galactan and oligofructose are used in cosmetic products. Dextran has found applications in human and veterinary medicine as blood plasma extender or blood flow improving agent. Furthermore it is used as cholesterol lowering agent.

Oligosaccharides gained interest in the food industry as they can be used as low cariogenic sugar substitutes (Crittenden and Playne, 1996), like trehalulose and isomaltulose.

Scientific and pharmaceutical interest also arose as many human pathogens initiate disease by utilizing their microbial adhesion proteins to attach to glycol-conjugates on host mucosal surfaces. Oligosaccharides of identical or similar structure to these naturally occurring ligands can both prevent bacterial attachment as well as mediate the release of attached bacteria (Johnson, 1999).

Another example of current efforts by the pharma industry to develop and design novel bioengineered oligosaccharides is the synthesis of erythropoietin (Epo), a glycosylated cytokine involved in the control of the red blood cell population, with the exact human glycosylation pattern (Sinclair *et al.*, 2005; Kim *et al.*, 2008).

1.4 Classification of carbohydrate active enzymes

The established system to classify an enzyme is the classification by the IUB-MB enzyme nomenclature, based on the type of reaction that enzymes catalyze and on their substrate-specificity. According to the first three digits of the E.C. classification, glycoside hydrolases (E.C. 3.2.1.x) are hydrolyzing *O*-glycosyl linkages. The last number indicates the substrate. However, for glycoside hydrolases a classification primarily based on the

substrate is too narrow and often ambiguous due to a broad specificity of the enzymes with often several substrates. To reflect also structural features, Henrissat (1991) proposed a new classification system for enzymes that build and breakdown complex carbohydrates and glycol-conjugates based on amino acid sequence similarity. Because there is a direct relationship between sequence and folding similarities, this classification correlates with enzyme mechanisms and protein folds more than enzyme specificity. Presently, the database comprises 118 families. Fructan- and glucan-forming enzymes belong to glycoside hydrolase (GH) families 68 and 70, respectively, which include enzymes responsible for the hydrolysis and/or transglycosylation of glycosidic bonds. All these enzymes use sucrose as their preferential donor substrate (Lammens *et al*., 2008; Cantarel *et al*., 2009). The free energy associated with the glycosidic bond in sucrose is 29 kJ/mol. This is sufficient to provide 13.5 kJ/mol required to sequentially add fructose (or glucose) units to a growing EPS chain (Schwarz *et al*., 2007).

1.5 Bacterial fructansucrases

Bacterial fructansucrases are extracellular enzymes capable of synthesizing levan and inulin from sucrose. According to the CAZy database they belong to the glycoside hydrolases family 68 (Cantarel *et al*., 2009). Depending on their products they are named levansucrase (E.C. 2.4.1.10) or inulosucrase (E.C. 2.4.1.9), respectively. Inulosucrases are exclusively found in lactic acid bacteria (LAB). In contrast, levansucrases are present in Gram positive as well as in Gram negative bacteria. At their amino acid level levansucrases from Gram positive and Gram negative bacteria share only a relatively low identity to each other of about 20 % (van Hijum, 2004).

Reactions catalyzed by fructansucrases

In general, fructansucrases are able to catalyze three different reactions depending on the acceptor of the transferred sugar unit. The enzymes cleave the glycosidic bond of their substrate sucrose (and in some cases raffinose) and use the energy released to catalyze the transfer of a fructosyl unit from sucrose to water (hydrolysis) (Figure 1-5) or to a growing fructan chain (transfructosylation) (Figure 1-6). In the presence of

suitable acceptors, fructansucrases synthesize fructo-oligosaccharides (oligosaccharide formation). Bacterial levansucrases catalyze all three reactions but with different efficiency.

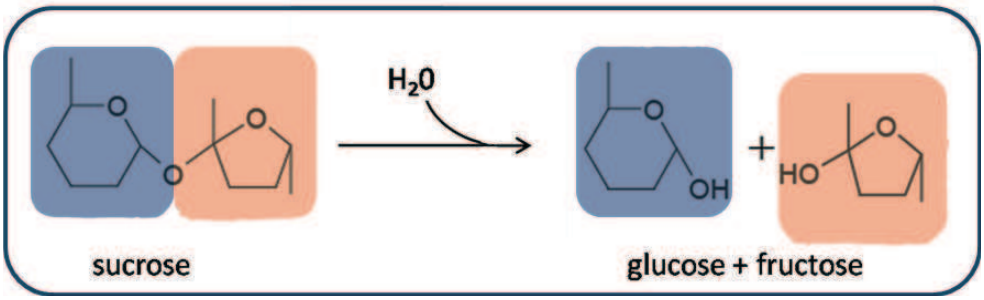

Figure 1-5: Schematic overview of the hydrolysis reaction catalyzed by bacterial levansucrases.

Fructan synthesis

Fructan synthesis starts with a transfructosylation reaction in which a sucrose molecule plays the role of fructosyl donor with a second sucrose molecule as the acceptor of the fructosyl moiety. Therefore bacterial fructans always possess a non-reducing glucose unit at the end of the polymer chain (French, 1993; van Hijum 2004). The first step in levan chain elongation is the formation of the trisaccharide 6-kestose (Chambert and Petit-Glatron, 1991). In the presence of low concentrations of sucrose (< 50 mM) known levansucrases display only hydrolytic activity with water as acceptor (Chambert and Petit-Glatron, 1991; Meng and Fütterer, 2003). At higher substrate concentrations, levansucrases catalyze the synthesis of high-molecular-mass fructan of the levan type (Figure 1-6) (Daguer *et al*., 2004).

It was shown that levansucrase SacB from *B. subtilis* synthesizes these high-molecular-mass fructan without transient accumulation of oligofructan molecules (Hernandez *et al*., 1995). This suggests a processive type of reaction for the fructan chain elongation where the growing polymer chain remains bound to the enzyme (Ozimek *et al*., 2006).

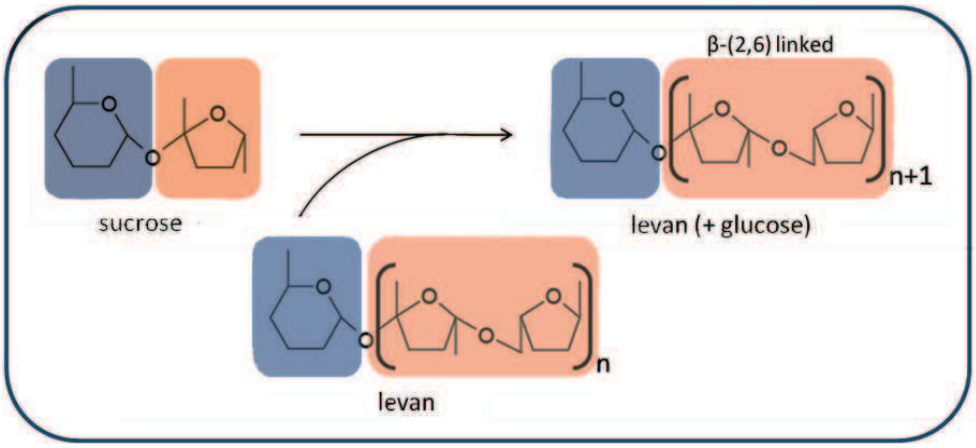

Figure 1-6: Schematic overview of the polymerization reaction catalyzed by a bacterial fructansucrases using the example of levansucrases. The release of the remaining glucose moiety is a byproduct of the reaction.

Acceptor reaction

In addition to the production of fructose polymers, levansucrases are also known for the synthesis of fructo-oligosaccharides which are produced by the transfer of fructose from sucrose (or raffinose) to suitable acceptors others than levan (van Hijum, 2004). Examples of acceptors to be considered are sucrose and kestose, yielding in the β-(2,1) linked oligosaccharides 1-kestose and nystose, respectively (Figure 1-7). Other acceptors include short-chain acylalcohols, sorbitol and various mono- to tetrasaccharides when present in high concentrations (Feingold *et al*., 1956; Hestrin *et al*., 1956; Hestrin and Avigad, 1958; Chambert and Petit-Glatron, 1991; Tieking, 2005).

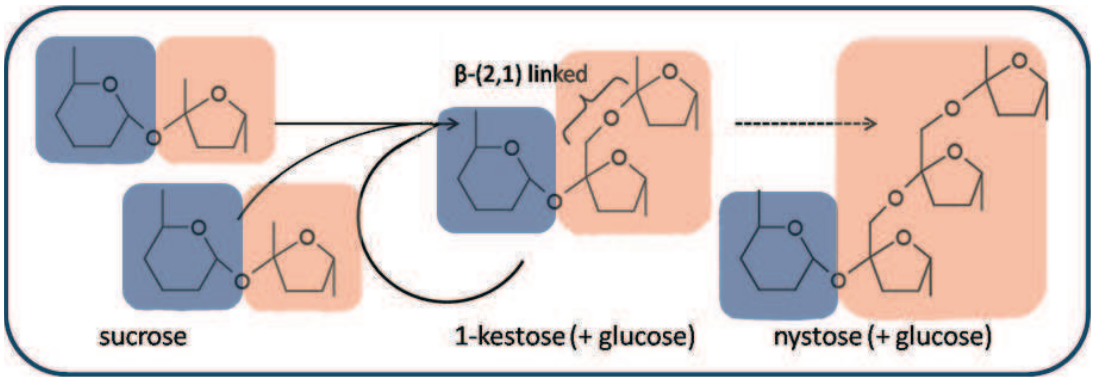

Figure 1-7: Schematic overview of the acceptor reaction catalyzed by bacterial levansucrases. The release of the remaining glucose moiety is a byproduct of the reaction.

Catalytic reaction mechanism

Glycoside hydrolase family 68 enzymes are retaining enzymes (Koshland and Stein, 1954). They cleave the glycosidic bond between subsite -1 and +1 (nomenclature according to Davies *et al*., 1997) (Figure 1-8) via a double displacement mechanism. The catalytic site consists of a conserved glutamic acid which acts as a proton donor for protonation of the glycosidic bond and an aspartic acid which performs the nucleophilic attack. A third acidic residue functions as a transition state stabilizer. The hydrolysis occurs in two steps involving a covalent fructosyl-enzyme intermediate and an overall retention of the anomeric configuration (Chambert *et al*., 1976; Rye *et al*., 2000). However, only limited information is available about the transfer of the fructosyl unit to an acceptor.

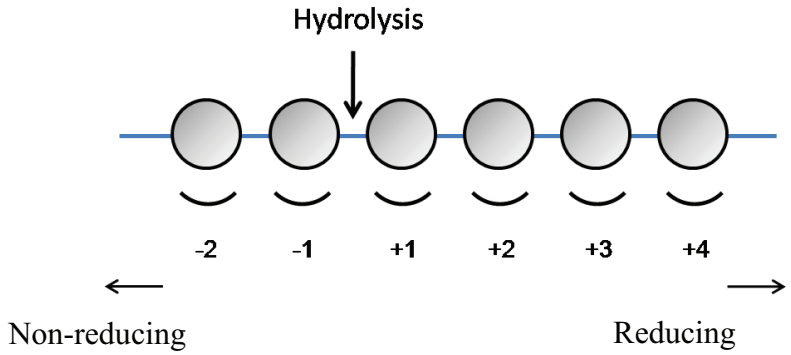

Figure 1-8: Diagrammatic representation of sugar binding site in fructansucrases (based on the scheme proposed by Davies *et al*., 1997.)

Structural information about GH 68 fructansucrases

Currently, high resolution crystal structures of two distinct levansucrases have been published (Meng and Fütterer, 2003, Martinez-Fleites *et al*., 2005) (Figure 1-9). However, the first reported crystal structure of a GH 68 family member was the structure of *Bacillus subtilis* levansucrase SacB, solved in 1980 at a resolution of 3.8 Å (LeBrun and van Rapenbusch, 1980). Today, coordinates of this structure are not available. However, interestingly the structure displayed a two domain protein with an elongated shape. Later, in 2003, Meng and Fütterer solved the high resolution structure (1.5 Å) of this enzyme (PDB code 1oyg). This structure shows a globular single domain protein, which suggests that the first structure reported did not give an accurate render of the levansucrase. The

high resolution structure of SacB from *B. subtilis* was followed by the one of LsdA from *Gluconacetobacter diazotrophicus* (PDB code 1w18) (Martinez-Fleites *et al*., 2005). These two enzymes share a five-bladed β-propeller fold enclosing the active site (Figure 1-9 A & C). Each β-sheet consists of four antiparallel β-strands with the classical "W" - topology. This fold is shared by the members of the Clan GH-J which harbors enzymes from GH 32 and GH 68 family. The deep negatively charged central pocket is composed mostly of residues belonging to the highly conserved sequence motifs, including invariant residues D86/D247/E342 and D135/D309/E401, respectively (Figure 1-9, B & D), which are also referred to as "catalytic triad" (Meng and Fütterer, 2003; Lammens *et al*., 2008).

Several site-directed mutagenesis experiments combined with biochemical analyses have been performed to identify key residues in substrate binding and recognition (Meng and Fütterer, 2003; 2008; Ortiz-Soto *et al*., 2008). Furthermore, two ligand-bound structures are available from *B. subtilis* levansucrase: a sucrose- and a raffinose-bound complex with the mutated enzyme (E342A) (PDB code 1pt2 and 3byn; Meng and Fütterer, 2003 and 2008). On the basis of these complex structures, the -1 and +1 subsite could be revealed by the unambiguous position of the glucose- and fructose in the active site. Cleavage of the glycosidic bond occurs between subsite -1 and +1. The complex structures revealed that the -1 subsite is highly specific for the sucrose moiety, whereas the +1 subsite shows more variability and allows binding of glucose (sucrose or raffinose as donor substrate) or fructose (sucrose or fructan as acceptor substrate) (Lammens *et al*., 2008).

The recently published structure of *Aspergillus japanicus* inulosucrase in complex with its natural acceptor substrates 1-kestose (PDB code 3ldr) and 1-nystose (PDB code 3lem) revealed further amino acid residues involved in the transfructosylation process (Chuankhayan *et al*., 2010). However, it still remains unclear if structural elements on the enzyme`s surface take part in the transfructosylation process and which residues remote from the active site may have impact on the enzyme´s product specificity.

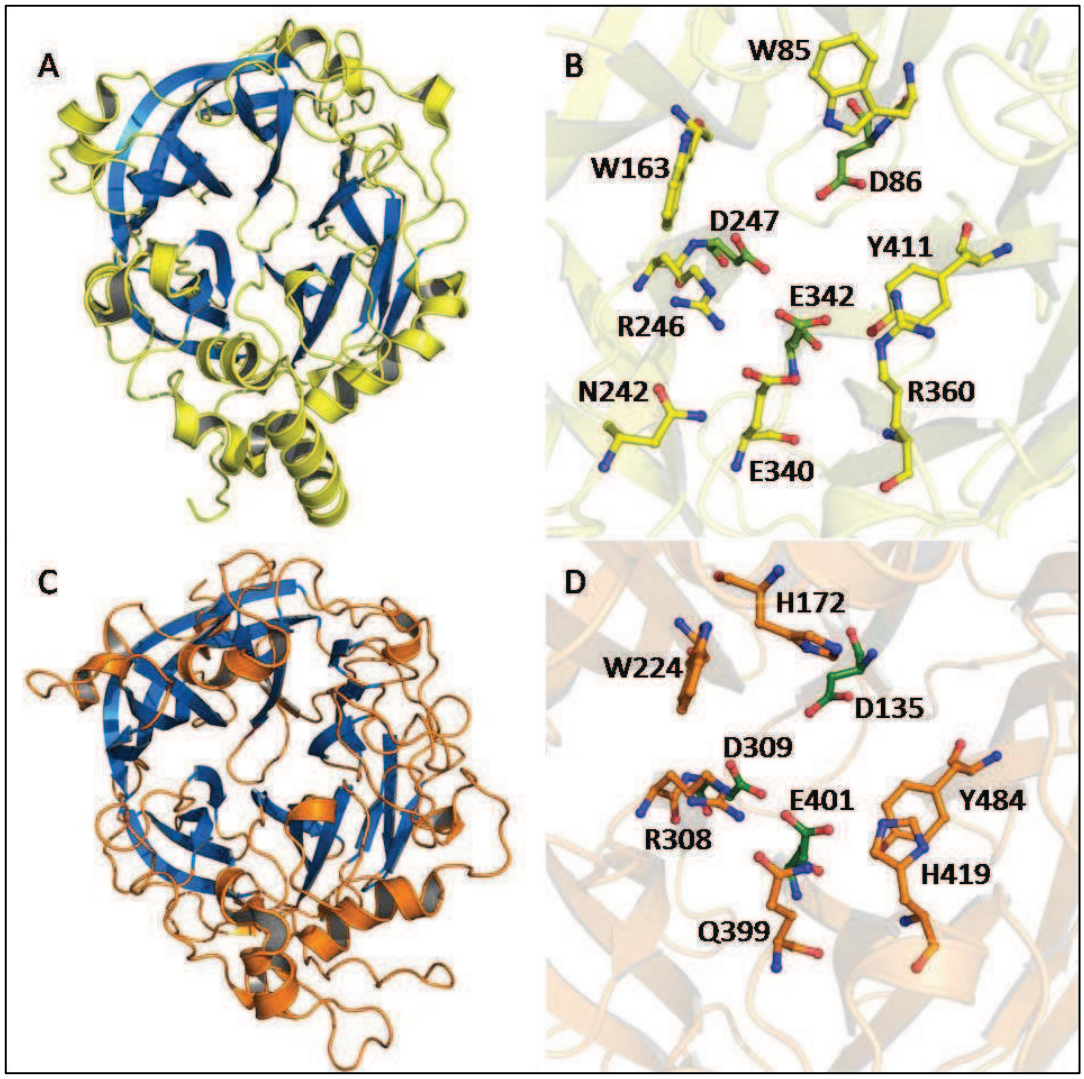

Figure 1-9: Crystal structures of GH 68 fructansucrases. The overall structure and the active site are presented. The three catalytic residues are depicted in green. *Bacillus subtilis* levansucrase (PDB code 1oyg) (A) and its active site (B); *Gluconacetobacter diazotrophicus* levansucrase (Martinez-Fleites *et al*., 2005) (C) and its active site (D).

Levansucrase SacB from Bacillus megaterium

The main subject of this thesis is on levansucrase SacB from *B. megaterium* DSM319 which is an extracellular enzyme with a relative molecular mass of 52.25 kDa. It was first identified by Dr. Rebekka Biedendieck from the group of Prof. Dr. Dieter Jahn (TU Braunschweig). SacB consists of 484 amino acids including a preceded N-terminal signal peptide of 29 amino acids length (Figure 1-10). On amino acid level SacB from *B. megaterium* shows an identity of 74 % to SacB from *B. subtilis*, why the protein from *B. megaterium* was also named SacB. It belongs to the glycoside hydrolase family 68, according to CAZy.

Figure 1-10: Schematic representation of the polypeptide chain of SacB from *B. megaterium*. The signal peptide is coloured in red, the catalytic domain in orange.

1.6 Bacterial glucansucrases

Bacterial glucansucrases (E.C. 2.4.5.1) of the GH 70 family are extracellular enzymes secreted by numerous bacteria e.g. by *Streptococcus* species from the oral flora, by the soil bacterium *Leuconostoc mesenteroides* and by lactic bacteria *Lactococci* (Henrissat and Davies, 1997; Monchois *et al*., 1999). To date, 108 enzymes have been classified in the GH 70 family. They are relatively large enzymes with an average molecular mass of 160 kDa, synthesizing exclusively α-glucan polymers (van Hijum *et al*., 2004). Based on the type of glucan that they produce, four distinct glucansucrases can be distinguished: dextransucrases, mutansucrases, alternansucrases and reuteransucrases (Crescenzi, 1995; Monchois *et al*., 1999).

It is known that glucansucrases produced by oral streptococci, play a key role in carcinogenesis process, as the synthesized glucan enhances the attachment and colonization of cariogenic bacteria like *Streptococcus mutans* and *Streptococcus sobrinus* (Hamada and Slade, 1980; Loesche, 1986).

Reactions catalyzed by glucansucrases

As bacterial fructansucrases, glucansucrases are also capable of catalyzing three different reactions depending on the nature of the acceptor. Glucansucrases cleave the glycosidic bond of their substrate sucrose and transfer a glucose unit either to water (hydrolysis), to a growing glucan chain (transglucosylation) or to a suitable acceptor (acceptor-reaction) (Monchois *et al*., 1999). In general, the mechanism of these reactions is only poorly understood. Structural information on these enzymes is still limited, but is essential for a deeper understanding.

Glucan synthesis

The synthesis of glucan is divided into three different steps: (1) initiation (2) elongation and (3) termination (Tsuchiya, 1953; Monchois *et al*., 1999). At the initiation of the glucan synthesis glucansucrases can be further classified into primer-dependent and primer-independent, depending upon whether they require a preformed glucan molecule for sucrose polymerization or not. However, the role of the primer remains controversial (Walker, 1978; Kingston *et al*., 2002). The glucan-chain-elongation is also poorly understood. There is still no unambiguous data confirming that the addition of glucose either occurs at the reducing end or at the non-reducing end. The termination of the glucan synthesis corresponds with the dissociation of the glucan from the enzyme (Monchois *et al*., 1999).

Acceptor reaction

In the presence of acceptors different from water and glucan, glucansucrases are able to produce low molecular weight oligosaccharides and α-glyco-conjugates by glycosylation of these acceptor molecules (Kraji, 2004). Suitable acceptors are the disaccharides maltose and isomaltose (Monchois, 1999; van Hijum *et al*., 2006), alkyl-glucosides (Lovell *et al*., 2003), molecules such as salicin, salicyl alcohol, phenol (Seo *et al*., 2005) or catechol, 4-methylcatechol and 3-methoxycatechol (Meulenbeld and Hartmans, 2000).

Catalytic reaction mechanism

Due to the various glucan structures and diverse oligosaccharides glucansucrases are able to produce and to utilize of a range of non-carbohydrate substrates. The catalytic mechanism of these enzymes is rather complex complicated and still not fully understood. Based on sequence similarity, a reaction mechanism similar to that of enzymes of the GH13 α-amylase family has been proposed for GH70 glucansucrases (van Hijum, 2004; Vujicic-Zagar, 2007). The reaction is supposed to proceed via a double displacement mechanism, involving a β-glucosyl-enzyme intermediate, retaining the α-anomeric configuration of the substrate in the product (Moulis *et al*., 2006; van Hijum *et al*., 2006; Vujicic-Zagar, 2007). The second part of the reaction, the transfer of the covalently linked

glucosyl unit to an acceptor remains largely unclear and is in contrast to the mechanism of GH13 enzymes still not clearly elucidated.

Structural and functional organization of glucansucrases

Although the reactions performed by glucansucrases and fructansucrases are very similar with respect to the use of sucrose as substrate, the enzymes neither share any sequence similarity nor a comparable overall structure (van Hijum, 2004). To date, there is no structural data available on a GH 70 glucansucrase in the protein data base (www.rcsb.org). However, recently a PhD thesis with limited access to some chapters was published with the structural characterization of the N-terminal truncated glucansucrase GTF180 from *Lactobacillus reuteri* 180 (residues 746 -1772). The crystal structure was solved at a resolution of 1.7 Å (Vujicic-Zagar, 2007). In contrast to previous assumptions that glucansucrases are composed of four distinct structural domains (Vickermann *et al*., 1996; van Hijum, 2004) the structure of GTF180-ΔN displays five domains, whereas the existence of two more domains is assumed at the lacking N-terminus. The five domains, designated as A, B, C, IV and V, show no linear arrangement along the polypeptide chain, but a "U-shaped" fold, in which each domain comprises one polypeptide segment coming from the N-terminal part of the peptide chain and one coming from the C-terminal part (forming the two stems of the "U") (Figure 1-11) (Vujicic-Zagar, 2007; B. Dijkstra, University of Groningen, pers. communication). Sequence alignments of GTF180 with other GH 70 enzymes show high similarity of between 30 and 80 %, what suggests a common fold for all GH 70 family members.

Comparison of the GTF180-ΔN domain arrangement to the ones proposed for other GH70 glucansucrases (Monchois *et al*., 1999) shows that the proposed "catalytic domain" corresponds to the N-terminal parts of domains B, IV and V. The putative "glucan binding domain" coincidences with the C-terminal part of domains B, IV and V (Vujicic-Zagar, 2007).

The five domains of GTF180-ΔN build an enzyme with an elongated shape. Domain C forms the bottom of the "U" and has an antiparallel β-sandwich-like fold. The catalytic core consists of domain A, which comprises a $(\beta/\alpha)_8$-barrel, and domain B which is folded into a 5-stranded twisted β-sheet. Domain IV is mostly α-helical and has no homology to

other proteins in the pdb, whereas domain V, which is thought to interact with carbohydrate components, is composed of series of repeats which fold into β-hairpins (Vujicic-Zagar, 2007; B. Dijkstra, University of Groningen, pers. communication).

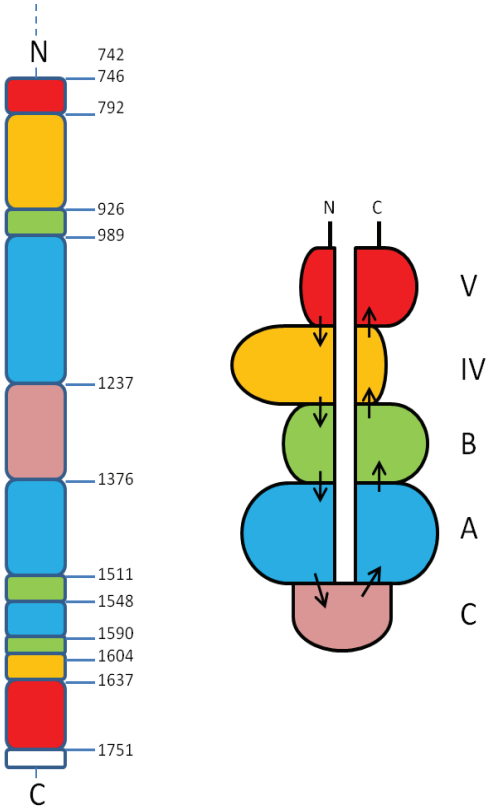

Figure 1-11: Schematic presentation of the "U-fold" course of the polypeptide chain of GTF180. Domains A, B, C, IV, and V are coloured in blue, green, purple, orange, and red, respectively (B. Dijkstra, University of Groningen, pers. communication).

The glucan binding domain of bacterial glucansucrases

Glucansucrases show a high affinity to α-glucans produced by themselves. Several studies demonstrated that the C-terminal domain of these enzymes is responsible for the binding of glucan. Therefore, it was also named glucan binding domain (GBD) (Kato and Kuramitsu, 1991; Lis *et al*, 1995; Kingston *et al*., 2002). In *Lactobacillus reuteri* GTF180-ΔN the glucan binding domain (domain V) comprises the first 47 and the last 114 residues. In general, GBDs are characterized by specific repeats in their primary structure. Giffard and Jacques (1994) analyzed these repeats and observed a common fundamental repeat: the YG-repeat with a length of 21 residues, which possible is the result of gene duplication events. The type and number of these YG repeats vary among the members of the GH70 family (Suwannarangsee *et al*., 2007). They are characterized

by the presence of clusters of aromatic residues and the occurrence of glycine from three to five amino acids downstream these clusters (Suwannarangsee *et al*., 2007). The clustered aromatic residues (tyrosine, tryptophan and phenylalanine) are supposed to stabilize the binding between sugar and protein by interacting with the sugar chain. Polar and acid residues may allow the formation of hydrogenbonds with hydroxyl residues of the polysaccharide (Quiocho, 1986; Monchois *et al*., 1999).

The precise function of the GBD as well as the interacting with glucan is largely unknown. The presence of the GBD seems to be necessary for the activity of the enzyme. Furthermore it was shown that the GBD is involved in glucan structure determination (Vickermann *et al*., 1996; Monchois *et al*., 1998; Monchois *et al*., 1999; Fujiwara *et al*., 2000). The significance of GBDs to the biology of the oral streptococci is often evaluated in the context of dental caries. It has been shown that glucans synthesized by glucansucrases make a significant contribution to the extracellular polysaccharide matrix which forms an integral component of the dental plaque biofilms (Banas and Vickerman, 2003). These glucans are exploited by bacteria, using glucan binding proteins to colonize smooth tooth surfaces.

Dextransucrase GtfR from Streptoccocus oralis

Glycosyltransferase R (GtfR) from *Streptococcus oralis* ATCC 10557 is a dextransucrase of 175 kDa (E.C. 2.4.1.5), belonging to GH 70 family. It is not much known about this enzyme. In the year 2000 the enzyme was characterized by Fujiwara and co-workers. Besides the production of dextran with α-(1,6) glucosidic linkages, GtfR catalyzes transglucosylation of different acceptors substrates, including alcohols and amino acids (Seibel *et al*., 2006; Hellmuth *et al*., 2007; Cantarel *et al*., 2009). It was shown that GtfR synthesizes from sucrose mainly dextran (86%). Glucose, leucrose and isomaltulose are just byproducts (Seibel *et al*., 2006). Investigations of Fujiwara and co-workers (2000) suggest that *S. oralis* GtfR strongly contributes to the establishment of oral bacteria biofilms.

2 AIMS AND SCOPE

This thesis was part of the Collaborative Research Centre SFB 578 (Sonderforschungsbereich 578) „Development of Biotechnological Processes by Integrating Genetic and Engineering Methods – From Gene to Product.“ Institutes from the Technical University of Braunschweig, the Helmholtz-Centre for Infection Research and the University of Applied Sciences of Magdeburg aimed in collaboration the combination and integration of methods from basic sciences and the fundamental engineering sciences in order to find a systematic approach for processes of protein production in recombinant microorganisms. One major aim of this collaboration was the investigation of a complete biotechnological process “From Gene to Product”, beginning with the first cloning step to the final downstream processing step of the target product. Therefore, the SFB was divided into four project areas (Hempel, 2006):

Molecular Biology of Product Formation
Systems Biotechnology of Product Formation
Process Technology
Application Technology

Each project area consists of two to seven subprojects. The outlined thesis was incorporated into subproject A7 “Structural Biology of Glycosyltransferases for the Optimization of Biotechnical Processes”. Aim of this subproject was the structural and functional investigation of the fructosyltransferase SacB from *Bacillus megaterium* and of the glucosyltransferase GtfR from *Streptococcus oralis*.

Bacterial fructansucrases and glucansucrases are able to synthesize numerous oligo- and polysaccharides. These target structures gained interest in the pharmaceutical as well as in the food industry. However, numerous questions regarding the catalytic mechanism of fructan- and glucansucrases remained to be answered. Therefore, the general aim of this thesis was to gain insight into molecular and biochemical processes underlying the synthesis of oligo- and polymers in *B. megaterium* SacB and *S. oralis* GtfR, respectively.

In order to investigate and gain a deeper understanding of the enzyme´s reaction mechanism, determination of three-dimensional crystal structures of enzymes in complex with substrate and products are aimed. These studies shall be the basis for rational design approaches in order to create enzyme variants with altered product spectra.

SacB from B. megaterium

The major part of this thesis included (i) establishing of a protocol for expression, purification and crystallization of SacB (ii) identification of functionally important amino acids and (iii) determination of the three-dimensional structure of SacB by X-ray analysis of SacB and variants crystals. Furthermore soaking and co-crystallization experiments with substrates, products and inhibitors were planned to gain structural information on SacB when complexed with a ligand.

The cloning, the biochemical characterization and product analyses of SacB were performed in cooperation with the group of Prof. Dr. Jürgen Seibel (University of Würzburg, Institute of Organic Chemistry) and the group of Prof. Dr. Dieter Jahn (Technical University of Braunschweig, Department of Microbiology).

GtfR from S. oralis

The aim of the second part of the thesis was the structural investigation of the C-terminal glucan-binding domain (GBD) of GtfR from *S. oralis*. To investigate this domain by X-ray crystallography, expression constructs had to be generated and protocols for the expression, purification and crystallization had to be established. A sufficiently high resolution crystal structure of a complex structure with different mono- and oligosaccharides sould provide detailed insights into structural features involved in glucan binding.

3 MATERIALS AND METHODS

3.1 Standard materials

If not stated otherwise, all chemicals used in this work were of '*pro analysis'* grade and were purchased from the companies Amersham Biosciences, Difco, Fluka, Hampton Research, Merck, Millipore, Qiagen, Riedel de Haen, Roche, Roth, Sigma, and Stratagene. Molecular biological methods used in this work were derived from standard collections of methods and protocols (Sambrook and Russell, 2000; Coligan, 2003; Ausubel *et al.*, 2007). These methods will not be explained in detail. Only essential modifications of standard protocols have therefore been described below.

3.1.1 Enzymes and molecular weight standards

Table 3-1: Enzymes.

Enzyme	Source
Benzonase	Merck
TEV-protease	HZI
PreScission® Protease	GE Healthcare
Thrombin Protease	Amersham Biosciences
SAP (Shrimp Alkaline Phosphatase)	Roche
T4 DNA Ligase	New England Biolabs
Platinum *Pfx* DNA Polymerase	Invitrogen
Restriction endonucleases: EcoRI, NcoI, BamHI, NdeI, NotI	New England Biolabs

Table 3-2: Molecular Weight Standards.

Name	Usage	Company
Benchmark	Protein: SDS-PAGE	Invitrogen
Fermentas Precision Plus Protein Standard All Blue	Protein: SDS-PAGE/Western Blot	Fermentas BioRad
Smart Ladder	DNA: Agarose gel electrophoresis	Eurogenetec

3.1.2 Crystallization screens

The following commercial screens were used to screen for initial crystallization conditions:

Table 3-3: Crystallization screens.

Screen	Company
The $AmSO_4$	Qiagen
The Anions	
The Cation	
The Classics I + II	
The Classics Lite	
The Cryo	
The MbClass I + II	
The Sparse Matrix 1-5	
The MPDs	
The PEGs	
The pHClear I + II	
SFP (JCSG+)	
Pro Complex	
PACT	
Grid Screen $AmSO_4$	Hampton Research Corp.
Grid Screen PEG/LiCl	
Crystal Screen Cryo	
Crystal Screen I + II	
Additive screen	
Cryo I + II	Jena Bioscience
Wizard I +II	

3.1.3 Plasmids

The following plasmids were used throughout.

pET28c (HZI)

Genes cloned into the *pET28c* vector are expressed with an N-terminal His_6-tag. A thrombin-protease cleaving site between tag and protein allows the target protein to be cleaved, releasing the protein of interest.

pETM30 and pQE60 (HZI)

Genes cloned into the vectors *pETM30* and *pQE60* are expressed with an N-terminal His_6-tag and glutathione S-transferase (GST) –tag. A TEV-protease cleavage site between tag and target protein allows the fusion protein to be cleaved, releasing the protein of

interest. These vectors were used to express the *gbd* of GtfR from *S. oralis*.

pGEX-6p-1 (GE Healthcare)

This vector allows the expression of the target protein fused to an N-terminal glutathione S-transferase (GST) –tag. A PreScission®-protease cleavage site between tag and target protein allows the fusion protein to be cleaved, releasing the protein of interest. The vector was used to express the *gbd* of GtfR from *S. oralis*.

pET11d (Rebekka Biedendieck, TU Braunschweig)

This vector was used by Dr. Rebekka Biedendieck for the generation of all *sacB*-constructs used in this work.

3.1.4 Bacterial strains

Table 3-4 summarizes the bacterial strains used in this thesis.

Table 3-4: Bacterial strains.

***E. coli* strain**	**Genotype**	**Source**
TOP 10	F^-, *mcr* A Δ(*mrr-hsd*RMS-*mcr*BC) φ80 *lac*Z Δ M15 Δ*lac*X74 *rec*A1 *ara*D139 Δ(*ara-leu*) 7697 *gal* U *gal* K *rps*L *end*A1*nup*G	Nitrogen
TG2	*supE hsdΔ5 thiΔ (lac-proAB) Δ(srl-recA) (306 :: Tn10) (tet^r) F′ [traD36proAB$^+$lacI^{q}lacZΔM15]*	Stratagene
BL21CodonPlus	B $F^–$ *ompT hsdS*($r_B^-m_B^-$) *dcm*+ Tet^r *gal* (DE3) *endA*	
(DE3)-RIL	Hte [*argU ileY leuW* Cam^r]	Nitrogen
Tuner™	F^-*omp*T *hsd*S_B ($r_B^-m_B^-$) *dcm*$^+$ *dam*$^-$ *gal-lac*Y1 (DE3)	Novagen

3.2 Buffers and Media

Media were sterilized by autoclaving (121 °C, 2 bar, 20 min, Infection Control, PST 9-6-9 HS1, BeliMed). Heat-sensitive additives were sterile filtered (pore width 0.22 µm).

Luria Bertani-medium (LB-medium):	Tryptone (Bacto)	10 g/L
	Yeast extract (Bacto)	5 g/L
	NaCl	5 g/L

SOB-medium:	Tryptone (Bacto)	20	g/L
	Yeast extract (Bacto)	5	g/L
	NaCl	10	g/L
	KCl	2.5	g/L
SOC-medium:	SOB-Medium + Glucose	2	mM
	$MgCl_2$	10	mM
	$MgSO_4$	10	mM

3.2.1 Antibiotics

Solid and liquid media were supplemented with antibiotics to ensure their selectivity. Antibiotics were added after media had cooled to below 50°C. Depending on the plasmid and bacterial strain, the following concentrations of antibiotics were used:

Table 3-5: Antibiotics.

Antibiotic	Final concentration	
Ampicillin	100	µg/mL
Chloramphenicol	34	µg/mL
Kanamycin	30	µg/mL

3.3 Microbiology

3.3.1 Agar plates

Agar plates for solid-phase cultivation of bacteria were made by adding 16 g agar to 1 L LB-medium. After autoclaving and cooling below 55°C, the appropriate antibiotics were added and the solution was poured into sterile petri dishes (Ø 10 cm) using a sterile clean bench. Kanamycin- and chloramphenicol-containing plates were stored for maximal six weeks at 4°C, ampicillin-containing plates for up to 2 weeks at 4°C.

3.3.2 Liquid cultures

Bacteria were grown in liquid cultures using Erlenmeyer flasks with 3-4 baffles to ensure sufficient aeration. The flasks, containing inoculated medium and appropriate antibiotics, were incubated at 20°C to 37°C and rotation speeds of 140 to 180 rpm. The rotation speed depended on the kind of culture (pre-culture, DNA-preparation, preparative or analytical expression).

3.3.3 Storage of bacteria

Bacterial strains were stored on agar-plates up to 10 days at 4°C. For long-term storage glycerol stocks were used. Therefore 5 mL of LB-medium were inoculated with a single clone from an agar plate and incubated for 14 h at 37°C and 180 rpm. Antibiotics were

added as required. A culture volume of 900 µL was mixed with 100 µL of sterile glycerol (87%). These stocks were frozen at -70°C.

3.4 Molecular biology

3.4.1 Preparation of competent cells

Electrocompetent cells

Electrocompetent cells were prepared using *E. coli* TOP10- or BL21 CodonPlus cells freshly plated on LB agar. A single colony was picked, transferred to 100 mL of LB-medium, and incubated o.n. (37°C, 180 rpm). 5 mL of this starter culture were transferred to 500 mL LB-medium. Bacteria were grown to an OD_{600} nm of 0.4, cooled rapidly and kept on ice for 30 min. The flasks and solutions subsequently used were sterilized and cooled to 4°C. Cells were centrifuged (4°C, 10 min, 5500 g), washed twice with ice-cold, sterile HEPES buffer (1 mM), once with ice-cold, sterile 10% (v/v) glycerol, and resuspended in a final volume of 1 mL 10% (v/v) glycerol (concentration $1\text{-}3 \bullet 10^{10}$ cells/mL). 60 µL aliquots were transferred into pre-cooled Eppendorf tubes and flash-frozen in liquid nitrogen. The electrocompetent cells were stored at -70°C.

Chemically competent cells

Chemically competent cells were prepared using *E. coli* Tuner, TOP10- or BL21 CodonPlus cells freshly plated on LB agar, respectively. A starter culture of 50 mL was inoculated from a single grown colony. After o.n. growth at 37°C and 180 rpm, 10 mL of the starter culture was used to inoculate the 500 mL of LB-medium. The cells were grown at 37°C to an OD_{600} of 0.5, then cooled and pelleted by centrifugation (4°C, 10 min, 5500 g). The cells were washed and resuspended twice in decreasing volumes of ice-cold and sterile solutions of 0.1 M $CaCl_2$. Finally cold glycerol was added to a concentration of 10% (v/v) and the cells were flashfrozen in liquid nitrogen and stored at -70°C in 60 µL aliquots.

3.4.2 Transformation of competent bacteria

For plasmid amplification or recombinant gene expression competent cells were transformed. By plating the bacteria onto LB-agar containing antibiotics for which resistance was conferred by the respective plasmid, positive transformants could be identified.

Electroporation

Electrocompetent *E. coli* cells were thawed on ice. 1-2 µL of plasmid DNA or ligation mixture were mixed with 60 µL of electrocompetent cells, transferred to a pre-chilled electroporation cuvette and incubated on ice for 1 min. The exterior wall of the cuvette was dried and placed into a Gene Pulser™ (BioRad). After an electric pulse of 2.5 kV, 200 Ω and 125 µF 1 mL of room-tempered SOC-medium was immediately added to the cells.

After incubating at 37°C with gentle shaking for 30 - 50 min, the transformants were plated on selective LB-agar and incubated o.n. at 37°C.

Heat shock transformation

1-2 µL of ligation mixture or plasmid DNA were mixed with 60 µL of chemically competent cell (thawed on ice) and incubated for 30 min on ice. A temperature shift (heat shock) to 42°C for 90 s was applied to the mixture. Followed by a rapid cooling on ice for 5 min. 800 mL of room-tempered SOC-medium were added to the cells followed by an incubation of 30 – 50 min at 37°C with gentle shaking. The transformants were plated on selective LB-agar.

3.4.3 Preparation of plasmid DNA

Depending on the amount of plasmid DNA required, plasmid DNA preparation from *E. coli* was carried out using either the Mini-kit or the Maxi-kit purchased from Qiagen. Therefore 5 mL (mini) or 500 mL (maxi) of LB-medium were inoculated with freshly transformed *E. coli* Tuner- or TOP10-cells. Bacteria were grown o.n. at 37°C and 180 rpm. The cells were harvested by centrifugation (5,000 g, 5 min, 4°C). Following steps were performed according to manufacturer´s instructions.

3.4.4 Determination of DNA concentration and purity

To determine the yield and purity of isolated bacterial plasmid-DNA, samples were measured spectrophotometrically. The absorption at 260 nm ($A_{260\ nm}$) was determined allowing the calculation of the DNA concentration, assuming that $A_{260\ nm}$ = 1.0 is equivalent to 50 µg/mL DNA. The ration of $A_{260\ nm}/_{280\ nm}$ was used to check the sample for protein contaminations. Ratios of 1.8 to 2.0 were considered as sufficiently pure to use the sample for subsequent cloning and transformation.

3.4.5 Agarose gel electrophoresis

Agarose gel electrophoresis is a method for analytical separation of DNA fragments according to their length. Therefore agarose gels consisting of 1% agarose in TAE buffer were prepared by boiling in a microwave and subsequent adding of ethidium bromide. The DNA samples were mixed with 10 x DNA loading dye before injection into the sample pockets. Gels were run with 5 V/cm for 60 min in TAE buffer. After the run the separated fragments were detected under UV illumination (λ = 254 nm).

TAE-buffer:	Tris	40 mM
	Sodium Acetate	20 mM
	EDTA	1 mM
	adjusted to pH 8.2 with acetic acid	

10 x loading dye:	sucrose	70	% (w/v)
	Bromphenol Blue	0.25	% (w/v)
	EDTA	100	mM

3.4.6 Extraction of DNA from agarose gels

PCR products and linearized vectors were extracted from agarose gels using the "Gel Extraction Kit" (Qiagen).

3.4.7 Digestion of plasmid DNA with restriction endonucleases

Digestion of DNA with restriction endonucleases is a standard procedure required for cloning and analysis of DNA fragments. Reaction buffers, conditions and concentration of DNA were chosen according to the New England Biolabs technical reference catalogue.

3.4.8 Dephosphorylation of linearized plasmid DNA

Dephosphorylation of linearized plasmid DNA is a method employed to avoid relegation of the original vector without target insert. For 5´-dephosphorylation 25 ng of linearized vector were incubated for 10 min with 1 U of Shrimp Alkaline Phosphatase immediately after digestion at 37°C. Inactivation was reached by heating to 65°C for 20 min.

3.4.9 Ligation of DNA fragments

Ligation of DNA inserts into linearized 5´-dephosphorylated vectors with T4-ligase was performed by mixing 1 U of the enzyme with ATP containing ligation buffer, 25 ng of the vector insert and a 5-fold molar excess of insert DNA. The mixture was incubated for 1 h at room-temperature, followed by 23 h at 16°C. After ligation, 2 µL of the mixture were used to transform competent cells.

3.4.10 Amplification of DNA by Polymerase Chain Reaction (PCR)

The amplification of target genes from template vectors was carried out by polymerase chain reaction. This reaction allows the cloning into respective recipient vectors. Standard PCR mixtures of 20 µL contained:

Template DNA	0.25-0.5	µg
Forward Primer	15	pmol
Reverse Primer	15	pmol
Deoxyribonucleotides (dNTPs)	300	µM (each)
Polymerase buffer	10	% (v/v)
$MgSO_4$	1	mM
Platinum Pfx DNA Polymerase	1	U

PCR was performed by using a PCR thermocycler. Standard program for *Pfx* Polymerase:

Initial denaturation	2	min	95°C
Denaturation	15	s	95°C
Annealing	30	s	[T_m-5 to 10 K]
Elongation	3	min	68°C
Final Elongation	7	min	68°C

Steps 2-4 were cycled 35 times.

The reaction was performed in a gradient cycler to optimize the yield of the amplification. PCR products were purified by agarose extraction kit (Qiagen) or PCR purification kit (Qiagen).

3.4.11 Design and synthesis of deoxyribo- oligonucleotides

Deoxy oligonucleotides were used as primers for PCR reactions. They were purchased by the company Operon (desalted and HPLC purified). To amplify genes of interest for re-cloning, restriction sites compatible with the target vectors were incorporated into tailor-made overhangs of primers. Cloning strategies were checked using VECTOR NTI (Invitrogen). All oligonucleotide primers are listed in Table 3-6.

Table 3-6: Oligonucleotides used in PCR reactions.

Name	Length [bp]	Sequence (5´→ 3`)	Restriction site	Target vector
pBluGtf-BamHI-Fwd	42	GAGAGGATC|CATGGGAAATTGGATTACTTCGATAAACGAGGC	BamHI	pGEX-6-P1
pBluGtf-NotI-Rev	42	GAGAGCGGC|CGCTCAGTTTCTTCCGAAGTTTTGACCTCTACC	NotI	pGEX-6-P1
GtfR-pETM30-Fw	38	GAGAC|CATGGGAAATTGGTATTACTTCGATAAACGAGG	NcoI	pETM30
GtfR-pETM30-Rw	40	GAGAGGATC|CTTAGTTTCTTCCGAAGTTTTGACCTCTACC	BamHI	pETM30
GtfR-pETM30-Fw	38	GAGAC|CATGGGAAATTGGTATTACTTCGATAAACGAGG	NcoI	pQE60
GtfR-yg-pQE60-Rw	37	GAGAGGATC|CGTTTCTTCCGAAGTTTTGACCTCTACC	BamHI	pQE60
GtfR-yg-pET28c-Fwd	39	GAGACA|TATGGGAAATTGGTATTACTTCGATAAACGAGG	NdeI	pET28c
GtfR-pETM30-Rw	40	GAGAGGATC|CTTAGTTTCTTCCGAAGTTTTGACCTCTACC	BamHI	pET28c

3.4.12 DNA sequencing

To confirm the successful modification of DNA sequence determination of the respective DNA region was performed by the company GATC Biotech using the dye terminator method. The results were provided as ABI chromatogram files. These chromatograms were compared and analyzed using the program CONTIGEXPRESS from VECTOR NTI (Invitrogen).

3.5 Protein production

Test expression

In order to find the optimal conditions for the recombinant expression of newly generated gene constructs, small-scale test expressions were performed prior to preparative synthesis. Expression constructs were transformed into suitable *E. coli* BL21 CodonPlus cells. Liquid pre-cultures of 5 mL LB-medium with appropriate antibiotics were inoculated with a single *E. coli* colony and grown over night at 37°C and 180 rpm. 1 mL aliquots of the pre-cultures (OD_{600} = 0.15) were used to inoculate 50 mL of expression test culture (LB-medium supplemented with required antibiotic). These test cultures were incubated at 37°C. After reaching an OD_{600} = 0.6–0.8, the expression temperature was shifted towards the desired value (20, 30 and 37°C) and the recombinant gene expression was induced by adding IPTG (0.1, 0.5 and 1 mM). Samples were taken before induction and at different time points after starting recombinant expression. OD_{600} was determined. Samples were centrifuged (12000 g, 5 min) and for lysis the resulting pellets were resuspended in 100 µL Crack-buffer per OD unit and heated at 98°C for 10 min.

For analysis of insoluble produced protein the overnight samples were additionally treated with bug buster™ (Novagen). 100 µL Bug Buster per OD unit were added to the pellets to extract the soluble fraction. After centrifuging, the remaining insoluble proteinaceous material was resuspended in an equivalent volume of 8 M urea. Aliquots of soluble and insoluble fractions were analyzed by SDS-PAGE.

Crack buffer:	Tris/HCl pH 6.8	50	mM
	SDS	2	% (w/v)
	Glycerol	10	% (v/v)
	Bromphenol blue	0.1	% (w/v)
	DTT	100	mM

Recombinant protein synthesis

SacB and variants

For recombinant production of SacB and its variants, cells of the *E. coli* strain BL21

CodonPlus were freshly transformed with the expression vector. A single colony from an agar plate was used to inoculate a pre-culture of 50 mL LB-medium supplemented with ampicillin. The pre-culture was incubated at 37°C over night, and then served as inoculum for 2 L of LB-medium containing ampicillin. Main culture was grown at 37°C and 180 rpm to OD_{600} = 0.6-0.8 followed by a temperature shift towards 28°C. Gene expression was induced with 0.5 mM IPTG and extended to 24 h.

Glucan binding domain of GtfR

For recombinant production of the GBD of GtfR cells of the *E. coli* strain BL21 CodonPlus carrying the desired plasmid were taken from a glycerol stock to inoculate a 50 mL pre-culture containing LB-medium and kanamycin. After overnight incubation at 37°C, the pre-culture was used to inoculate 2 L of selective LB- medium. The main culture was grown at 37°C and 180 rpm to OD_{600} 0.6 – 0.8. After a temperature shift towards 25°C gene expression was induced by adding IPTG to a final concentration of 0.5 mM and extended to 4 h.

Harvest and lysis of E. coli cells

All steps were performed on ice or 4°C to prevent proteolytic degradation of the target proteins.

After expression, cells were harvested by centrifugation (6000 g, 15 min). The supernatant was discarded and the cell pellet was resuspended in 30 mL ice-cold 50 mM Sorensen's phosphate buffer (pH 6.9 for SacB and pH 6.0 for GBD) with 1 µL Benzonase® (4.2 U/µL) per liter of bacterial culture and protease inhibitor cocktail "Complete EDTA-free" (Roche). The cell suspension was disrupted either by 5 cycles of sonication at 50% power for 20 s each using a SonoPlus HD201 (Bandelin) or by high pressure cell disrupter (Constant Systems, 20 kPa, 2 runs). Soluble and insoluble components were separated by centrifugation at 37,000 g for 1 h.

Phosphate buffer after Sorensen

		pH 6.0	**pH 6.9**
KH_2PO_4	% (v/v)	65.3	52.3
Na_2HPO_4	% (v/v)	34.7	47.7

3.6 Protein purification

Affinity chromatography

Affinity chromatography is based on strong and specific binding properties of a so-called affinity tag to an immobilized ligand. Proteins with a terminal polyhistidine tag (His tag) bind via this affinity tag to Ni^{2+} nitrilotriacetic acid (Ni-NTA) under native conditions in

low- and high salt buffers.

The glucan binding domain of GtfR was produced with an N-terminal His-tag. Soluble cell lysate containing the recombinant fusion proteins was coupled with 4 mL of buffer-equilibrated Ni-NTA beads (Qiagen) for 1 h at 4°C. Immobilized proteins were washed with 300 mL wash-buffer (100 mM Tris buffer, pH 7.8). Stepwise addition of elution buffer (wash-buffer supplemented with 0.5 M imidazole; steps: 10, 50, 200 and 500 mM imidazole) resulted in elution of the bound proteins. The flow-through was checked for proteinaceous material with a flow-through UV monitor at 280 nm. The success of the purification was checked by SDS-PAGE. After elution and subsequent dialysis against buffer without imidazole, proteolytic cleavage was initiated by addition of 40 Units of thrombin and incubation o.n. at 4°C. By applying a second affinity chromatographic step the cleaved affinity-tag binds to the immobilized Ni-NTA while the target protein is found in the flow-through.

Ion Exchange Chromatography (IEC)

By ion exchange chromatography (IEC), protein mixtures were separated depending on their electrostatic properties. The first purification step of SacB and variants was adapted from a previous protocol of Meng and Fütterer (2003). An 80 mL CM-Sepharose™ column was equilibrated with at least 2 column volumes (CV) of 50 mM Sorensen's phosphate buffer, pH 6.9 (buffer A). The supernatant of cell lysis was applied on the column (flow rate 0.5 mL/min) and the column was washed with 1 CV of buffer A to remove unbound proteins (flow rate 0.8 mL/min). After binding and washing the retained proteins were eluted by a linear phosphate gradient using buffer A and 500 mM Sorensen's phosphate buffer, pH 6.9 (buffer B). The eluate was collected in 3 mL fractions. Fractions containing SacB were identified using SDS-PAGE, pooled and concentrated by ultracentrifugation.

Ceramic Hydroxyapatite (CHT) chromatography

CHT, a sintered material turned from crystalline to ceramic, possesses different functional groups such as Ca^{2+}, PO^{3-}_4 and OH groups. Negatively charged carboxyl groups of the protein can interact with hydroxyapatite calcium ions and are competed out by a phosphate gradient.

Gel permeation chromatography (GPC)

Gel permeation chromatography (GPC) utilizes the correlation of molecular retention times in a porous gel matrix and the physical size of the molecules. The gel matrix contains pores of a given range of diameters, causing a negative correlation between size of a molecule and its retention time on the column. All GPC experiments were performed on ÄKTA FPLC Systems (GE Healthcare) with a HighLoad Superdex 75 (16/60) column. Prior to the GPC the protein solutions containing SacB were concentrated to 1-3 mL and filtered (pore width 0.22 µm). SacB solution was applied onto the column equilibrated with 50 mM NaCl, 10 mM Mes, pH 6.0. Proteins were separated with 1.5 CV of buffer (~180 mL) with a flow rate of 0.5 mL/min and collected in fractions of 1.5 mL.

3.6.1 Concentrating protein solutions

Protein solutions were concentrated by ultracentrifugation using Vivaspin-2, -6 and -20 concentrators with appropriate MW cut-off. The pore size was chosen to be at least 20 kDa smaller than the size of the target protein. The absorption of the concentrate was checked at 280 nm against the flow-through until the desired concentration was obtained.

3.7 Protein biochemical methods

3.7.1 Determining protein concentrations

The concentration of a given protein was determined via the measurement at A_{280} nm against buffer on the basis of the Beer-Lambert law. The molar extinction coefficient ε_{280} (Gill and von Hippel, 1989) for each protein or fragment was calculated *in silico* with the program VectorNTI (Invitrogen) or the PROTEINCALCULATORTOOL (www.scripps.edu/~cdputnam/protcalc.html). Since this method does not distinguish proteins in a mixture, it was only used when the target protein had largely been purified. Alternatively, the Nanodrop spectrophotometer system (PEQLAB Biotechnology) was used where only 2 µL of a given protein solution needs to be applied for photometrical determination of a protein concentration. The estimated extinction coefficients for SacB and the GBD of GtfR are:

Protein construct	**Extinction coefficient ε_{280}**
B. megaterium SacB variants	60450 M^{-1} cm^{-1}
S. oralis His_6- GBD	82640 M^{-1} cm^{-1}

3.7.2 SDS-Polyacrylamide gel electrophoresis (SDS-PAGE)

Proteinaceous samples were analyzed by SDS-PAGE (Laemmli, 1970), a type of electrophoresis used to separate proteins by size. Samples of 2 – 10 µL volume were mixed with 2 µL of SDS-containing 8 x loading buffer and brought to a final volume of 12 µL with water. Heating of the protein samples to 95°C for 10 min ensured complete unfolding. After a short centrifugation step, the samples were loaded onto the stacking gel. The polyacrylamide concentration of the gels was chosen according to expected protein sizes, varying between 12 and 15% (v/v). Gels were run at 40 mA per gel for 40 - 60 min.

Table 3-7: Composition of SDS-polyacrylamide gels.

Solution	Separating Gel		Stacking Gel
	12%	15%	5%
Acrylamide/Bisacrylamide 30%(w/v)/ 0.8% (w/v)	12 mL	15 mL	1.5 mL
4 x Lower buffer	7.6 mL	5 mL	
4 x Upper buffer			2.5 mL
10% SDS	0.3 mL	0.2 mL	
H_2O	10 mL	4.7 mL	5.9 mL
TEMED	40 µL	20 µL	15 µL
25% APS	100 µL	50 µL	25 µL

8 x Sample buffer:	10% SDS	16	mL
	Glycerol	4	mL
	Tris/HCl, pH 6.8	2.2	mL
	β-mercaptoethanol	800	µL
	Bromphenol blue	1	spatula tip
4 x Upper buffer:	Tris/HCl, pH 6.8	0.5	M
	SDS	0.4	% (w/v)
4 x Lower buffer:	Tris/HCl, pH 8.8	1.5	M
10 x SDS-PAGE running buffer:	Tris Base	30.2	g
	Glycine	144.1	g
	SDS	10	g
	add H_20	1	L
Staining solution:	Ethanol	30	% (v/v)
	Acetic acid	10	% (v/v)
	Coomassie R-250	0.25	% (w/v)
Destaining solution:	Ethanol	40	% (v/v)
	Acetic acid	10	% (v/v)
Dehydrating solution:	Methanol	50	% (v/v)
	Ethanol	3	% (v/v)
	H_2O	47	% (v/v)

3.7.3 Western blotting

After separation of proteins by SDS-PAGE, protein samples were transferred from gels onto a polyvinylidene difluoride (PVDF) membrane for subsequent immunodetection or N-terminal sequencing experiments. Freshly run gels were equilibrated for 15 min in transfer buffer, together with 4 pieces of gel-sized Whatman paper. The PVDF membrane was activated in 100% methanol for 15 min and thereafter equilibrated for 1 min in transfer buffer. For the transfer a semi-dry blot apparatus was used. The gel was placed onto the membrane and sandwiched between two layers of soaked Whatman paper onto the anode of the apparatus. Proteins were blotted onto the membrane for 20 min at 18 V per gel. The success of the procedure was checked via reversible staining of the membrane with Ponceau Red. Samples intended for N-terminal sequencing were cut out of Ponceau Red stained membrane band with a scalpel.

Transfer buffer:	CAPS	10 %
	Methanol	10 %
	H_20	80 %
Poinceau Red staining solution:	Ponceau S	0.2 %
	Acetic acid	2 %

Immunodetection of His-tagged proteins

His_6-tagged fusion proteins were detected using a specific antibody against the His-tag of the fusion protein. After the transfer (see section 3.7.3), the membrane was “blocked” by incubation for 1 h at 4°C with 5% (w/v) skim milk powder in PBS-T (PBS supplemented with 0.1% (v/v) TWEEN-20) to prevent non-specific binding of the antibody. The saturated membrane was then washed 5 times for 5 min with PBS-T and incubated with the antibody for at least 1 h at RT. After antibody binding, the membrane was washed again 3 times for 5 min. The peroxidase bound antibody was visualized on the membrane using an enhanced chemiluminescens (ECL) kit (Amersham Bioscience).

3.7.4 N-terminal sequencing

Protein bands separated by SDS-PAGE, blotted onto PVDF-membrane (see section 3.7.4) and visualized by Ponceau Red staining were N-terminal sequenced by automated Edman degradation (Edman and Begg, 1967). All sequencing experiments were carried out by Rita Getzlaff (HZI, Braunschweig) using an ABI 494A HT Protein Sequencer (Applied Biosystems).

3.7.5 Dynamic light scattering

Dynamic light scattering (DLS) allows the determination of the particle size distribution of a protein solution. As a rule of thumb, a polydispersity of the sample of 20% was used as a good indicator for a monodisperse protein suitable for crystallization. The experiments were carried out on a DynaPro 801TC instrument (ProteinSolutions/Wyatt Technologies).

Protein samples (5 – 10 mg/mL) were centrifuged at 21,000 x g for 15 min at 4°C prior to the DLS experiment. 15 µL of the protein solution were introduced into the cell and the scattering intensity fluctuations were recorded at ambient temperature until the signal stabilized. The data were analyzed assuming a globular shape of the proteins using the instrument software DYNAMICS (ProteinSolutions).

3.7.6 Mass spectrometry

All mass spectrometry experiments in this work were carried out by the biophysical analytics group of Dr. Manfred Nimtz (HZI, Braunschweig) with an Ultraflex TOF/TOF mass spectrometer (Bruker). Salt was removed from the protein samples by dialysis against ddH_2O. If protein bands were cut out of SDS gels, the samples were washed three times with ddH_2O to remove residual SDS.

3.7.7 Thermal shift assay

The fluorescence-based thermal shift assay is a method to check the relative stability of proteins in the presence of different buffers, additives, and cofactors ((Niesen *et al*., 2007). It uses the hydrophobic fluorescent dye Sypro® Orange which is added to the protein solution. The protein solution is heated in decent steps and the fluorescence is measured. Upon undergoing thermal denaturing and unfolding of the protein the dye binds preferentially to the now exposed hydrophobic residues, and the increase of the fluorescence signal is recorded as a function of temperature. The melting points of proteins in different conditions are determined. The condition leading to the highest melting point is most suitable to stabilize the protein and thus increases probably the crystallization process.

The experiments were set up in 96-well plates. Thermal stability was measured in a real-time PCR machine (BioRad) where the plate was heated from 20°C to 95°C in 0.5°C steps. Each well contained 40 µL of an individual buffer, 5 µL of protein in concentrations between 0.5 and 2.0 mg/mL and 5 µL of Sypro® Orange Data processing was done with the software BioRad MyIQ5.

3.8 Protein crystallization

3.8.1 Initial screening

Initial screening for crystallization conditions was carried out in 96-well plates, using

different commercially available screens (Nextal/Qiagen, Hampton Research and Jena Bioscience, see also section 3.1.3) and sitting drop vapor diffusion at 20°C and 4°C. 200 – 500 nL drops composed of equal amounts of protein- and crystallization solution were pipetted using a nano-drop pipetting robot (Mosquito, TTP LabTech). Reservoir solutions were pipetted in a volume of 70 nL. The plates were sealed with Manco™ Crystal Clear tape (Jena Bioscience) and incubated at 20°C and 4°C.

3.8.2 Optimization

Initial crystallization hits were optimized manually in 24-well sitting-drop and hanging-drop plates using drop volumes of 2-3 µL and varying the salt/precipitant and protein concentration as well as the pH value, ionic strength and temperature to produce crystals suitable for X-ray diffraction experiments.

Seeding techniques

Seeding is a method to introduce pre-formed crystal nuclei into a drop to control nucleation and alter the way in which crystals grow in order to obtain large, single high quality crystals. In this study streak seeding, macroseeding and microseeding was employed, whereas the latter method was the most successful. Microseeding involves the introduction a few tiny fragments of crystalline matter are introduced in crystallization solutions with lower supersaturation than it is required for spontaneous nucleation (Rupp, 2009). Therefore, an existing crystal was transferred to 1.5 mL reaction tube together with 50 µL reservoir solution and thoroughly vortexed for 10 min to crush the crystal into pieces. A dilution series of crushed crystal fragments diluted with equal volumes of reservoir solution was prepared and used in new crystallization setups. Crystallization drops were pipetted using 1 µL of protein solution plus 1 µL of 2^6, 2^8, 2^{10}, 2^{12}-fold dilution.

3.8.3 Co-crystallization and soaking

In this study a range of potential binding partners for SacB was used for co-crystallization and soaking experiments. Aliquots of 6-kestose and allosucrose were provided by the group of Prof. Dr. Jürgen Seibel (University of Würzburg) in the solid state.

Co-crystallization

Several attempts were undertaken to co-crystallize SacB and its variants with substrate, products or a potential inhibitor. At this, protein and binding partner were mixed and incubated before the crystallization was set up. The final concentration of the ligand was chosen to be in a 2- to 10- fold molar excess to the protein to achieve reasonably binding site occupancy. Incubation lasted in different setups between 30 min and 1 day, either at RT or at 4°C. To remove excess ligand after incubation, the protein-ligand mixture was either transferred into dialysis devices and dialyzed against the appropriate buffer without ligand or applied on a desalting column and washed with buffer.

Crystal soaking

Crystal soaking of substrate, products and an inhibitor has been applied in the case of SacB variants. Soaking experiments were performed either with solutions of the respective soaking compound or with the compound used as a solid. Soaking solutions contained the compound in different solutions (1, 10, 20, 50, 100 and 150 mM). The solutions were dissolved either in water, reservoir solution or in the buffer, in which the protein was present. To a 2 µL sitting drop experiment 0.2 µL of soaking solution was successively added. Solid substrate was added in small crumbles to the crystallization droplets containing the crystals. Soaking times varied between 5 min to 1 day.

3.8.4 Crystal transfer experiments

In order to remove bound PEG molecules from the active site of SacB variant crystals, a transfer of SacB crystals from the PEG 1000 containing mother liquor to solutions with another precipitant was performed. According to the protocol described by Schreuder and co-workers (1988) a series of solutions containing increasing concentrations of alternative precipitants were pipetted into the reservoirs of a hanging- drop set up. Drops of the original mother liquor were placed on the cover slips. Subsequently, drop solution and the reservoir solution are allowed to equilibrate for a few days via vapor diffusion. If the drop with the original mother liquor attracts water, then the concentration of the alternative precipitant is too low with respect to the original mother liquor. If the drop loses water then the concentration of the alternative precipitant is too high. It is assumed that, in order to protein crystals intact in different solvents, the solution with the different precipitant should neither attract water from the crystal nor donate water to the crystal. The changes in water content of the droplets were measured by determining the volume of the droplets with a pipette and by eye. The initial volume of the droplets was 10 µL.

3.9 Data collection, structure determination and refinement

Protein crystals larger than 30 µm in the smallest dimension were harvested from their mother liquor and flashed-cooled in liquid nitrogen. For cryo-protection of SacB crystals an aliquot of the original crystallization condition was mixed with glycerol to yield a final concentration of 25%.

The majority screening of crystals for their potential to diffract X-rays was performed in part at the HZI home source being composed of a rotating copper anode generator (RINT-2000 series RU-H3R) and an R-AXIS IV^{++} image plate detector (Rigaku). Further screening and data set collection was performed at synchrotron beamlines BW7A, X11, X12 and X13 (DESY, European Molecular Biology Laboratory, Hamburg, Germany), BL 14.1 and BL 14.2 (BESSY, Berlin, Germany) and ID 14-2, ID 23-1, ID 23-2, ID 29 and BM 16 (ESRF, Grenoble, France) using charge coupled device detectors (Mar CCD or Mar555 Flatpanel from MarResearch or ADSC).

X-ray diffraction data were indexed and integrated with the XDS package (Kabsch, 2010), HKL2000 (Otwinowski and Minor, 1997) or with Mosflm (Leslie, 1992) Images

were scaled using SCALA (ccp4 suite, 1994) or XScale (Kabsch, 2010). Data were further processed and manipulated using the CCP4 Suite of software (1994).

Structure determination

Molecular replacement is a phasing technique which can be used when a search model structurally similar to a molecule in the target crystal structure is available. The method of molecular replacement was applied to solve the structures of SacB D257A and N252A. SacB from *Bacillus subtilis* (PDB entry: 1OYG) served as the search model. The program MOLREP (Vagin and Teplyakov, 2010) was used for MR.

Difference Fourier

In the case of isomorphous crystals that only contain small structural changes caused by inhibitor binding, metal removal or replacement or a local mutation of one or several amino acids, the structure can be determined by the difference Fourier technique (Messerschmidt, 2007). Due to almost identical unit cell parameters the structures of SacB variants K373A, Y247A and Y247W were solved by using the method of difference Fourier with the structure of N252A as the starting model.

Twinning analysis

Analysis of data quality and twinning analysis in particular was performed using the programs PHENIX Xtriage (Zwart *et al*., 2005) and SFCHECK (Vaguine *et al*., 1999).

Model building and structure refinement

ARP/wARP (Lamzin and Wilson, 1993) was used to build first models into the initial electron densities. Examination and interpretation of the electron density maps were done using program COOT (Emsley *et al*., 2010). Structure refinement aims at improving the agreement of an atomic model with observed diffraction data and chemical restraints. The initial models were refined by alternating cycles of manual fitting and restraint refinement using REFMAC5 (Murshudov *et al*., 1997). The $2F_o$-F_c map was usually contoured at 1.0 σ and the F_o-F_c map was at 3.0 σ. Twin refinement was performed using the amplitude based twin refinement option in REFMAC5.

Water placement and ligand fitting

ARP/wARP Solvent or COOT was used to place water molecules into weighted (F_o-F_c) difference density maps. Difference density of at least 3.0 σ with a spherical shape and the presence of suitable donors/acceptors were chosen as criteria to place a water molecule into it. After refinement the newly introduced water molecules were inspected visually in difference maps ($2F_o$-F_c). Water molecules were deleted if they had *B* factors higher than 60 $Å^2$.

Small molecule ligands were fit manually into (F_o-F_c) difference density maps using

COOT (Emsley *et al*., 2010). Geometric parameters and topology files for bound PEG molecules were obtained from the HIC-UP server (Kleywegt *et al*., 2003). Stereochemical restraint libraries for ligands were directly obtained from REFMAC5 (Murshudov *et al*., 1999).

Structure validation and deposition

A number of validation and deposition tools were used for checking the geometry and stereochemistry of the structures.

MOLPROBITY

Molprobity is a general-purpose web service offering quality validation for 3D structures of proteins, nucleic acids and complexes (Chen *et al*., 2010).

PDB ADIT

The Auto Dep Input Tool (ADIT) was developed by the RCSB for the deposition and validation of structures to be included in the pdb archive (http://deposit-beta.rcsb.org/adit/).

SFCHECK

SFCHECK is a software package for evaluating the structure factor data obtained from X-ray diffraction experiments and for assessing the agreement between the atomic model and these data (Vaguine *et al*., 1999).

3.10 Figure preparation

All figures representing structures were prepared using the software PYMOL (DeLano, 2002) and ChemSketch (ACD/Labs). Sequences were aligned using ClustalW (http://www.ebi.ac.uk/Tools/clustalw2/index.html) and displayed using ESPript (Gouet *et al*., 1999).

4 RESULTS

4.A SacB from *Bacillus megaterium*

In the first part of this thesis the levansucrase SacB from *Bacillus megaterium* was investigated. The studies were performed in cooperation with Arne Homann and Sven Götze from the group of Prof. Dr. Jürgen Seibel (University of Würzburg, Institute of Organic Chemistry) and with Dr. Rebekka Biedendieck and Martin Gamer from the group of Prof. Dr. Dieter Jahn (Technical University of Braunschweig, Department of Microbiology).

4.A.1 Identification of the active site residues

For the identification of the active site residues of SacB from *B. megaterium*, the amino acid sequence of the mature protein was aligned against the sequences of levansucrases already known and biochemically characterized (Homann *et al*., 2007). Eleven amino acids in the active site were identified and subsequently substituted by alanine to eliminate the functional groups at the selected positions. To examine the effect of active site mutations on the enzymatic activity, the variants products were analyzed with thin layer chromatography and the reaction kinetics were characterized (Homann *et al*., 2007). The active site of GH 68 levansucrases is described as a catalytic triad (van Hijum, 2006). Based on sequence similarities, the catalytic triad of SacB from *B. megaterium* was suggested to be comprised of residues Asp95, Glu352 and Asp257. Substitution of these amino acids with alanine inhibited the enzyme activity (Homann *et al*., 2007). In addition, structural and kinetic information from the homologue levansucrase from *B. subtilis* showed that Asp95 acts as the nucleophile in the hydrolysis reaction. The corresponding acid/base catalyst is Glu352. Asp257 is supposed to coordinate the fructofuranosyl moiety in position 3-OH and 4-OH. In order to obtain insights into the reaction mechanism of SacB, the wildtype enzyme as well as the variants D257A and E352A were selected for crystallization and structure determination.

4.A.2 Identification of amino acids remote from in the active site of SacB with impact on the transfructosylation

By structure-based alignments of levansucrase SacB from *B. megaterium* with levansucrases SacB from *B. subtilis* and the related LsdA from *G. diazotrophicus* (Martinez-Fleites *et al*., 2005), eight additional amino acid residues with potential impact on the polysaccharide synthesis mechanism of SacB were identified. Asn252 which is also located outside the active site was identified in a random mutagenesis screen to have a strong impact on transfructosylation. An exchange to alanine abolishes polymer function (Beine *et al*., 2008). Based on these results Martin Gamer from the group of Prof. Dr. Dieter Jahn (Department of Microbiology, TU Braunschweig) generated plasmids encoding 13 SacB variants. All SacB variants were kinetically characterized by Arne Homann (2009). The substitution of Asn252 with alanine had a complete stop of the polysaccharide production as a consequence. Regarding the variant K373A, the formation of tri- and tetrasaccharides was lowered while the synthesis of penta- and hexasaccharides was slightly enhanced. The exchange of tyrosine in position 247 to alanine led to the formation of shorter oligosaccharides. Nona- and decasaccharides were enhanced and the transfructosylation stopped after approximately 9 fructosyl units. However, the exchange of tyrosine to tryptophan yielded the same oligosaccharide pattern as the wild-type SacB. The reason may be that an unpolar favoured π-π-stacking mechanism is possible with tyrosine as well as tryptophan but not with alanine. Due to their clear impact on the product specificity, from 13 characterized variants, N252A, K373A, Y247A and Y247W were chosen for crystallization and structure determination in order to extent our understanding of the molecular mechanism of the transfructosylation process.

4.A.3 Structure determination of SacB variants

4.A.3.1 Production and purification

To obtain SacB in amounts sufficient for biochemical characterization and crystallization experiments, it was produced in *E. coli* BL21 CodonPlus cells. After testing different expression conditions, optimal production was achieved with bacterial cells grown in LB

medium until reaching an optical density (OD_{600}) of 0.7. Typical yields amounted to 4 mg per L of culture. Following cell disruption by sonification, purification of overexpressed *sacB* was achieved by a combination of different chromatography techniques. Key biophysical data on SacB are summarized in Table 4-1. Initial purification of the ~52 kDa wildtype enzyme and its variants was performed by means of an ion exchange chromatography (IEC) step at pH 6.0 using a CM-sepharose column (Figure 4-1).

Table 4-1: Physio-chemical parameters of SacB variants, as calculated with VectorNTI (Invitrogen).

Parameter	SacB
amino acids	484
Molecular mass (kDa)	51.15
Molar extinction coefficient	60450
Isoelectric point (pI)	5.48

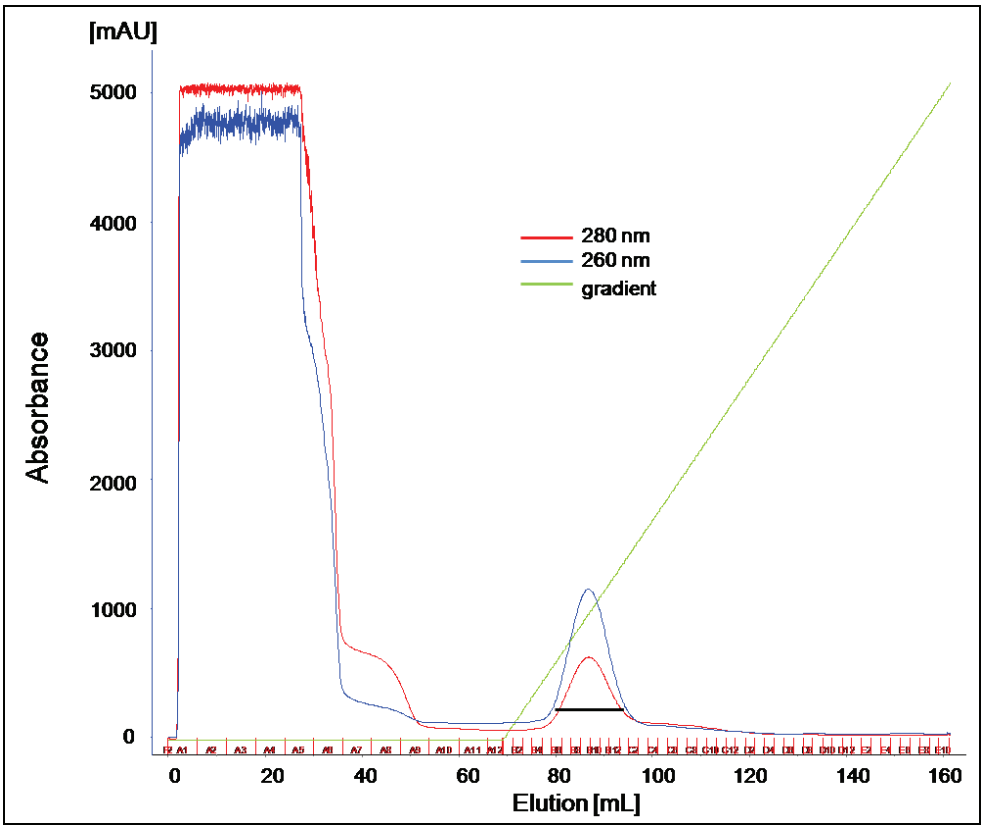

Figure 4-1: Chromatogram of SacB purification via Cation exchange chromatography.
(CM-Sepharose CL6B, 80 mL, GE Healthcare); Solid lines: absorption at 280 nm (blue), 260 nm (red); and applied phosphate gradient (green). SacB was almost completely separated from the crude protein extract. Fractions pooled and used for subsequent purification are marked by a black bar.

Although a very precipitous phosphate gradient was applied, the final elution fractions of this first purification step already displayed a very high degree of purity as demonstrated by SDS-PAGE (Figure 4-2).

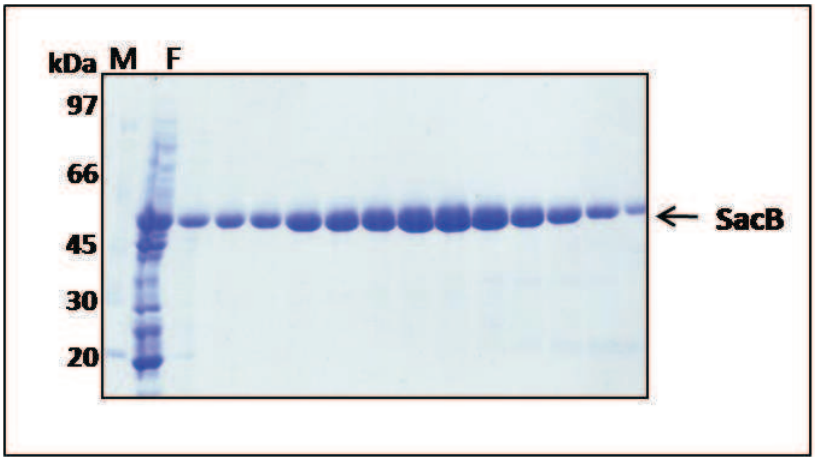

Figure 4-2: Purification of SacB via IEC. Coomassie brilliant blue stained SDS-PAGE: with unbound (F), and elution fractions from the cation exchange chromatography step, performed with *E. coli* crude extract containing SacB. The first lane contains the protein standard (M) with respective molecular masses indicated in kDa.

After cation exchange chromatography, ceramic hydroxyapatite (CHT) chromatography was chosen as an additional purification step. CHT chromatography separated several high and low molecular impurities from SacB. This purification step was additionally introduced after several unsuccessful crystallization trials, to improve the crystallization. Gel permeation chromatography was introduced as the last step of purification to remove aggregated protein and small molecular weight contaminants (Figure 4-3 and 4-4).

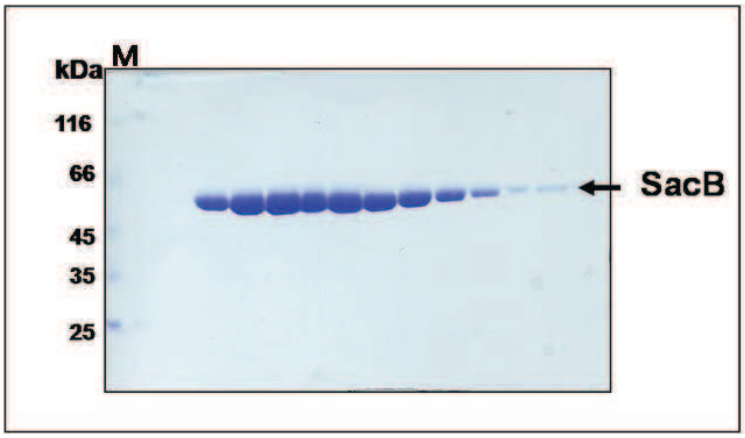

Figure 4-3: Purification of SacB via Gel permeation chromatography. Coomassie brilliant blue stained SDS-PAGE: Peak fractions of the GP chromatography. The first lane contains the protein standard (M) with respective molecular masses indicated in kDa.

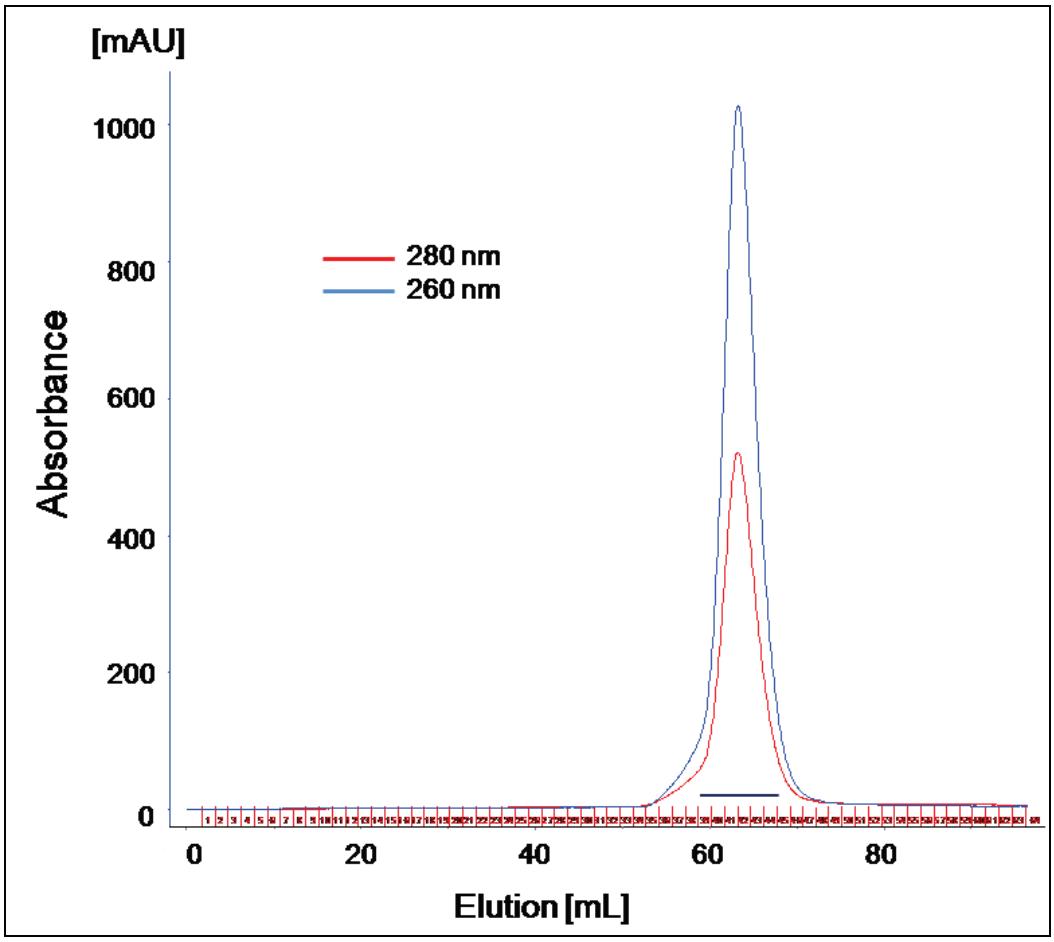

Figure 4-4: Chromatogram of SacB purification via Gel permeation chromatography using a HiLoad-16/60-Superdex-75 prep grade column. Solid lines: absorption at 280 nm (blue), 260 nm (red). Fractions collected and used for subsequent crystallization are marked with a black bar.

4.A.3.2 Crystallization of SacB

After purification the polydispersity of all protein samples was analyzed by DLS measurements. Monodisperse (polydispersity ≤15 %) and purified fractions of wildtype SacB and variants could easily be concentrated to final concentrations between 8 and 10 mg/mL. Nanoliter-scale screening for crystallization conditions, using commercial screening kits, resulted in two different conditions found suitable for crystallization. Small rod-shaped micro-crystals appeared after four weeks. Subsequent optimization of the initial conditions led to the growth of larger crystals at 20°C in 24-well hanging- and sitting drop vapor diffusion plates within one week. The first optimized condition was composed of 0.1 M Na-phosphate-citrate pH 4.1, 0.1 M $MgSO_4$ and 30% (v/v) PEG 400 (condition A), the second one was composed of 0.1 M Na-phosphate-citrate pH 4.1,

0.2 M lithiumsulfate, and 20 % (w/v) PEG 1000 with or without 1 mM $CaCl_2$ (condition B). Similar crystals were obtained under a variety of conditions ranging from pH 3.8 to 4.5 using Sorensen's phosphate buffer. PEG sizes from 500 to 4000 resulted in the same crystal habits. Although these optimized crystals diffracted X-rays to higher resolutions than the initial crystals, the maximum resolution of 3.5 - 4.0 Å still was not sufficient to solve the structure. Each SacB variant crystallized under both conditions but in different quality. The final crystals of SacB D257A were obtained by using condition A (Figure 4-6 A). Best crystals of variants N252A, K373A, Y247A and Y247W were obtained from condition B (Figure 4-6 D). Reproduction of crystals from condition A as well as from condition B turned out to be very difficult. In several crystallization experiments it was found that better results could be obtained by using freshly prepared protein. Due to the poor diffraction quality of SacB crystals, independent on the crystallization condition, different seeding methods were applied in order to optimize the crystal quality.

4.A.3.3 Improving crystallization by seeding experiments

Seeding is a method to introduce pre-formed crystal nuclei into a crystallization drop in order to control nucleation and alter the way in which crystals grow. It uses the basic conditions under which the crystals normally grow, sometimes slightly modified. In the case of SacB, macroseeding (Zhu *et al*., 2005), microseeding and streak-seeding (Bergfors, 2003) were used to obtain useful crystals. However, macroseeding did not result better diffracting crystals. Streak seeding was performed with a cat´s whisker (Stura and Wilson, 2003) or a horse tail hair swiping across a crystal and streaking through the new drop. Small crystals became visible after two days, but this method did not yield any crystals of suitable size for diffraction experiments (Figure 4-5).

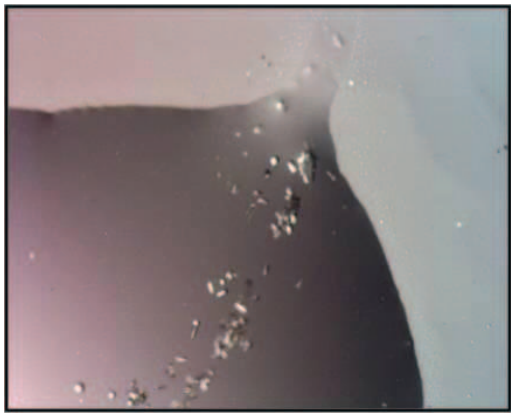

Figure 4-5: Streak seeding. Crystallization drops were set up with wild type SacB. Crystal nuclei have been placed into the crystallization drop using a cat´s whisker.

Final crystals of SacB D257A, suitable for X-ray experiments were obtained after microseeding experiments under condition A (Figure 4-6; B-C). Best crystals of variants N252A, K373A, Y247A and Y247W were obtained after microseeding experiments from condition B (Figure 4-6; E-F). In spite of successful microseeding experiments, the reproduction of high quality crystals from condition A as well from condition B still turned out to be very difficult and had to be repeated various times.

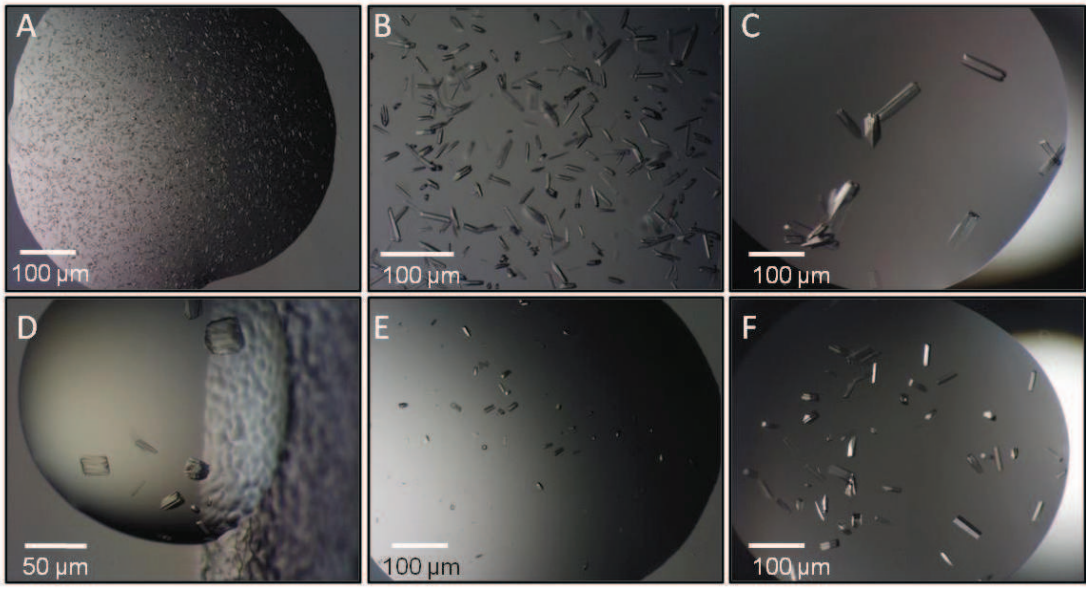

Figure 4-6: Crystallization of SacB variants.
(A-C) Crystals of SacB variant D257A (condition A): Initial screening for crystallization conditions resulted only in micro-crystalline precipitate (A); Optimization of these conditions yielded in larger crystals (B); and microseeding in crystals suitable for X-ray experiments (C).
(D-F) Crystals of variant Y247W (condition B): Small crystals after initial screening (D); after optimization of the condition (E) and larger crystals after performance of several cycles of microseeding (F).

The wild-type SacB and variant E352A represented an exception in crystallization behavior and yielded in spite of numerous micro- and streak-seeding trials neither larger nor well diffracting crystals. The obtained crystals were still of a small size (≤ 10 µm) and no structure of the wild-type enzyme and the variant E352A could be determined until the end of this thesis.

4.A.3.4 X-ray data collection and structure determination of SacB D257A

Optimized rod-shaped crystals of SacB variant D257A diffracted X-rays to 1.85 Å using synchrotron radiation. An X-ray diffraction data set of 380 images using rotation

increment of 0.3° per image was collected at beamline BM16 at the ESRF (Grenoble, France). Indexing X-ray diffraction data was achieved using the HKL2000 program package (Otwinowski and Minor, 1997). Analysis revealed that this variant crystallized in orthorhombic space group $P2_12_12_1$, with unit cell dimensions of a = 49.0, b = 55.3 and c = 163.6 Å. The calculation of Matthews probabilities (Matthews, 1968) based on the molecular mass of 51.15 kDa, suggested one molecule per asymmetric unit (V_M = 2.17 $Å^3$ Da^{-1}) with a solvent content in the crystals of 43.26 %.

The crystal structure of SacB D257A was determined by molecular replacement method using MOLREP with the coordinates of SacB from *B. subtilis* (PDB code: 1oyg, Meng and Fütterer, 2003) as the search model. The molecular replacement solution had a correlation coefficient of 58.4%. The F_o–F_c difference map showed continuous densities also for regions which were absent in the starting model, thus indicating a correct molecular replacement solution. The first cycle of refinement resulted in an initial R_{work} of 32.21 and an R_{free} of 35.08. Final structure model was yielded after iterative cycles of manual model building and refinement using COOT (Emsley *et al*., 2010) and REFMAC5 (Murshudov *et al*., 1999), respectively. In the last cycle water molecules were introduced to the model using *ARP*/*wARP solvent* (Lamzin and Wilson, 1993). Electron density was visible for SacB residues 34-381. For data collection and refinement statistics see Table 4-2 and 4-4, respectively.

4.A.3.5 X-ray data collection and structure determination of SacB variants N252A, K373A, Y247A and Y247W

Crystals of variants N252A, K373A, Y247A and Y247W grown from condition B, were harvested using 25 % (v/v) glycerol as cryoprotectant and employed in X-ray diffraction experiments. Diffraction data were collected at synchrotron beamlines BL 14.1 and BL 14.2 at the BESSY (Berlin, Germany). Synchrotron data were processed with XDS program package (Kabsch, 2010). Indexing of the diffraction images revealed that these crystals belonged to monoclinic space group $P2_1$, in contrast to crystals grown from condition A. The solvent content was estimated to be about 43 %, which corresponds to the presence of four molecules in the asymmetric unit.

Due to a different space group, the structures of variants N252A, K373A, Y247A and Y247W could not be solved by difference Fourier method using the structure of variant D257A. Thus, the model of SacB variant D257A was used to solve the structure of variant N252A by molecular replacement using MOLREP from the ccp4 Suite (1994). Afterwards, the structures of the remaining three variants could be solved by the method of difference Fourier due to almost identical unit cell dimensions, using REFMAC5 (Murshudov *et al*., 1999). For details of data processing and refinement see Table 4-2 and 4-4, respectively.

Table 4-2: Data collection statistics.

Diffraction data	**D257A**	**N252A**	**K373A**	**Y247A**	**Y247W**
	ESRF	BESSY	BESSY	BESSY	BESSY
Beamline	BM16	14.1	14.2	14.2	14.2
Wavelength (Å)	0.97881	0.91841	0.91841	0.91841	0.91841
Space group	$P2_12_12_1$	$P2_1$	$P2_1$	$P2_1$	$P2_1$
Resolution (Å)[a]	24.66 -1.90	47.59-2.00	47.77-1.75	47.77-2.00	47.72-1.90
	(2.00-1.90)	(2.11-2.00)	(1.84-1.75)	(2.11-2.00)	(2.00-1.90)
Unit cell	a = 49.0	a = 93.4	a = 93.6	a = 93.8	a = 93.5
dimensions	b = 55.3	b = 100.2	b = 100.1	b = 100.0	b = 100.1
	c = 163.6	c = 95.2	c = 95.5	c = 95.5	c = 95.5
	β = 90	β = 90.7	β = 90.7	β = 90.6	β = 90.5
Unique reflections[a]	35,974	117,898	174,958	118,542	135,669
	(5140)	(16,939)	(25,307)	(16,927)	(18,574)
Completeness (%)[a]	99.9 (100)	99.7 (98.6)	98.8 (97.9)	99.6 (97.9)	98.1 (92.4)
Multiplicity[a]	3.4 (2.6)	3.6 (3.0)	4.0 (4.0)	4.5 (3.2)	3.4 (2.6)
R_{merge} (%)[a]	21.1 (58.7)	10.0 (43.4)	10.7 (52.2)	22.5 (58.0)	12.6 (42.4)
I/σI [a]	7.5 (2.9)	9.4 (2.9)	10.50(2.6)	7.5 (2.4)	8.3 (3.8)
Wilson *B*-factor					
($Å^2$)	17.6	33.0	22.7	26.4	26.4
Molecules/AU	1	4	4	4	4
Solvent content (%)	42.3	42.6	42.9	42.9	42.8

[a] Values in parentheses refer to the highest resolution bin

For each variant four molecules were placed in the asymmetric unit without any clashes. Refinement cycles were alternated with manual model building using the program COOT (Emsley *et al*., 2010). Water molecules were automatically identified and added by the program *ARP/wARP solvent (*Lamzin and Wilson, 1993*).*

Twinning

During initial data processing twinning analysis was performed using the program PHENIX Xtriage (Zwart *et al*., 2005). Twinning is a physical phenomenon in crystals where at least two separate but symmetrically related lattices exist in the same crystal (Endo-Streeter, 2009). For data sets of variants N252A, Y247A and Y247W analysis revealed pseudo-merohedral twinning. This is a form of twinning in which the lattice has a higher symmetry than the symmetry of the unit cell content (Zwart *et al*., 2008). The twinned crystals had twinning fractions between 8 % and 16 %, indicating just partial twinning. The program REFMAC5 allowed refinement against those twinned data, using the "amplitude-based" twin refinement option. The twin fractions could be refined without any negative impact on the overall refinement process. However, the refinement statistics for untwinned data were slightly better than for twinned.

4.A.3.6 Secondary structure elements of SacB

Secondary structure elements in SacB variants were determined with DSSP (Kabsch and Sander, 1983) and STRIDE (Heinig and Frishman, 2004) with identical results for all five variants. The secondary structure assignment differed only minimal between the two programs. In Figure 4-7, the manually checked consensus assignment is shown together with the protein sequence.

4.A.3.7 Structural overview of SacB

The crystal structures of SacB variants D257A, N252A, K373A, Y247A and Y247W have been determined at reasonable resolutions of 1.9, 2.0, 1.75, 2.0 and 1.9 Å, respectively. Their refinement converged to relatively good crystallographic R-factors (16.1 – 20.1). All five structures presented here appear to be very similar. D257A (space group $P2_12_12_1$) contains one molecule in the asymmetric unit, whereas the other four

variants (space group $P2_1$) comprise four molecules in the asymmetric unit. Each of the four molecules in the asymmetric unit adopts a similar overall conformation and the monomers have a relatively low root mean square deviation with respect to each other (Table 4-3).

Table 4-3: Average r.m.s.d. for backbone atoms of the four monomers of a variant to each other.

SacB variant	Average r.m.s.d. (Å)
N252A	0.527
K373A	0.705
Y247A	0.391
Y247W	0.487

Comparison of the models exposed only rather minor conformational differences among the structures of the five variants; hence a representative structural overview is referred to variant D257A (unless specifically indicated).

The structured part for variant D257A as well as for the other four variants consists of residues 34 – 481. As expected, structurally, the enzyme shares the same overall architecture as its homologue SacB from *B. subtilis*. *B. megaterium* SacB folds into a single globular domain with a diameter of about 50.0 Å with a five-fold β-propeller motif as described for all known structures of the clan GH J enzymes. Each blade of the five-fold β-propeller (labeled I-V, Figure 4-8, A) consists of four antiparallel β-strands showing the classical `W` topology, labeled A-D.

The five propeller blades are arranged around a propeller axis, which runs roughly vertical in Figure 4-8 B. As in all β-propeller structures, the β-strands forming the blades, are strongly twisted (Alberto *et al*., 2004). The angle between the axis of the first and the last β-strand of a blade amounts approximately 90°. Short stretches of single 3_{10} helices can be observed between several individual β-strands. The N-terminus runs along the perimeter of blades IV and V and forms a clamp-like loop which adds a fifth β-strand (residues 44 - 46) to blade III. Before finally leading into blade I, two more short antiparallel β-strands can be found. The N-terminus is stabilized by a hydrogen bond from

residue Asn289 to Phe37. Asn289 belongs to helix $3_{10}6$, which is sharply kinked at this residue to facilitate stabilization by forming this hydrogen bond.

The central cavity of the β-propeller is formed by highly conserved sequence motifs of the GH family 68, including amino acids 93-VWD-95, 172-WSGS-175, 349-DEIER-353, 256-RDP-258, and 420-TYS-422. Each of these motifs belongs to the β-strand A of the particular blades.

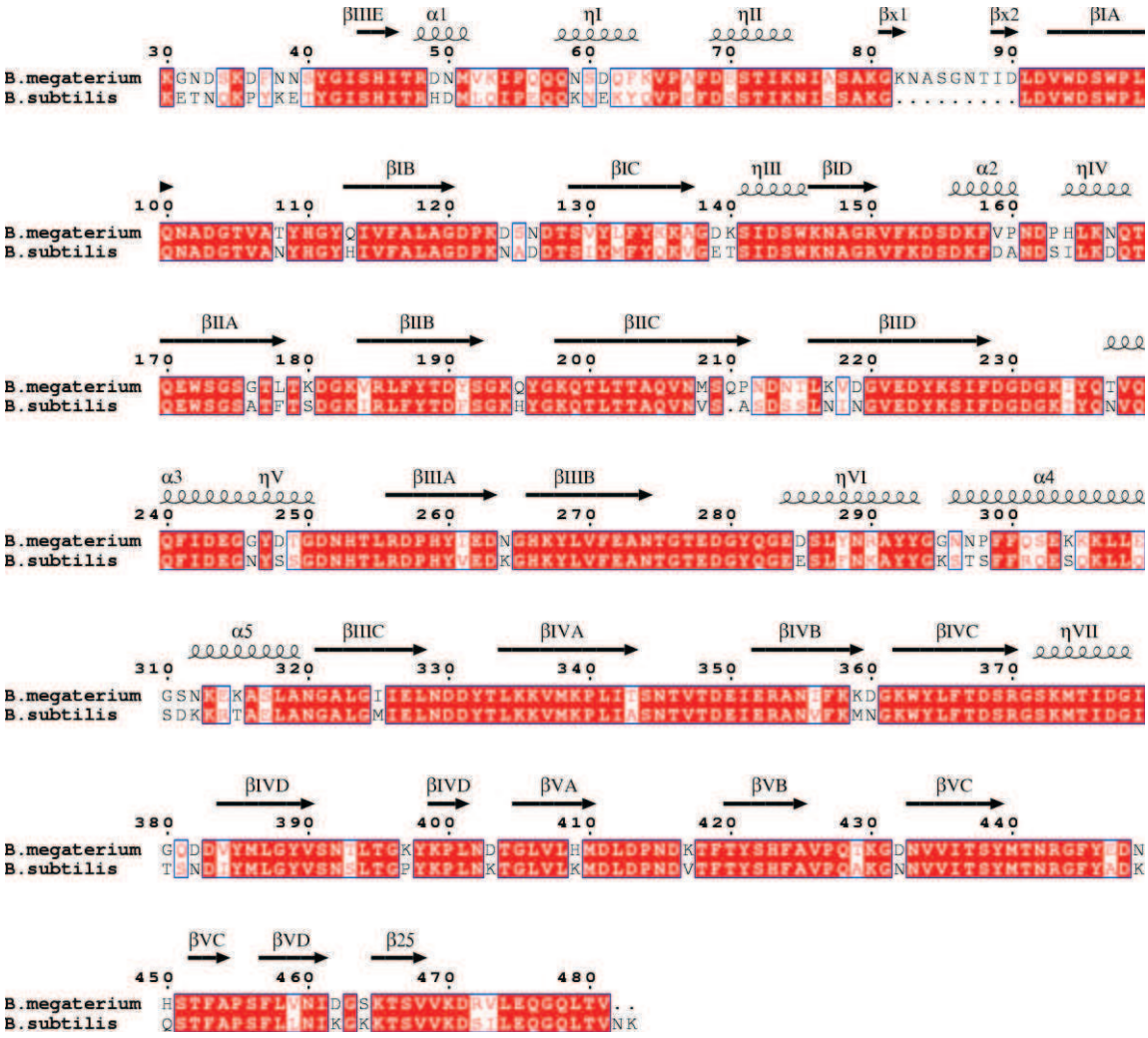

Figure 4-7: Sequence alignment of SacB from *B. megaterium* and SacB from *B. subtilis* and secondary structure assignment for SacB from *B. megaterium*.

The sequence alignment shows a sequence identity of 74 % between the two enzymes. The secondary structure assignment was performed with the ESPRIPT server (Gouet *et al.*, 1999). Secondary structural elements are indicated by spirals (helices and 3_{10} helices (η)) and arrows (β-strands) above the sequence. The β-strands are labeled by a Roman numeral, denoting β-strands I –V, and a capital letter, A-D.

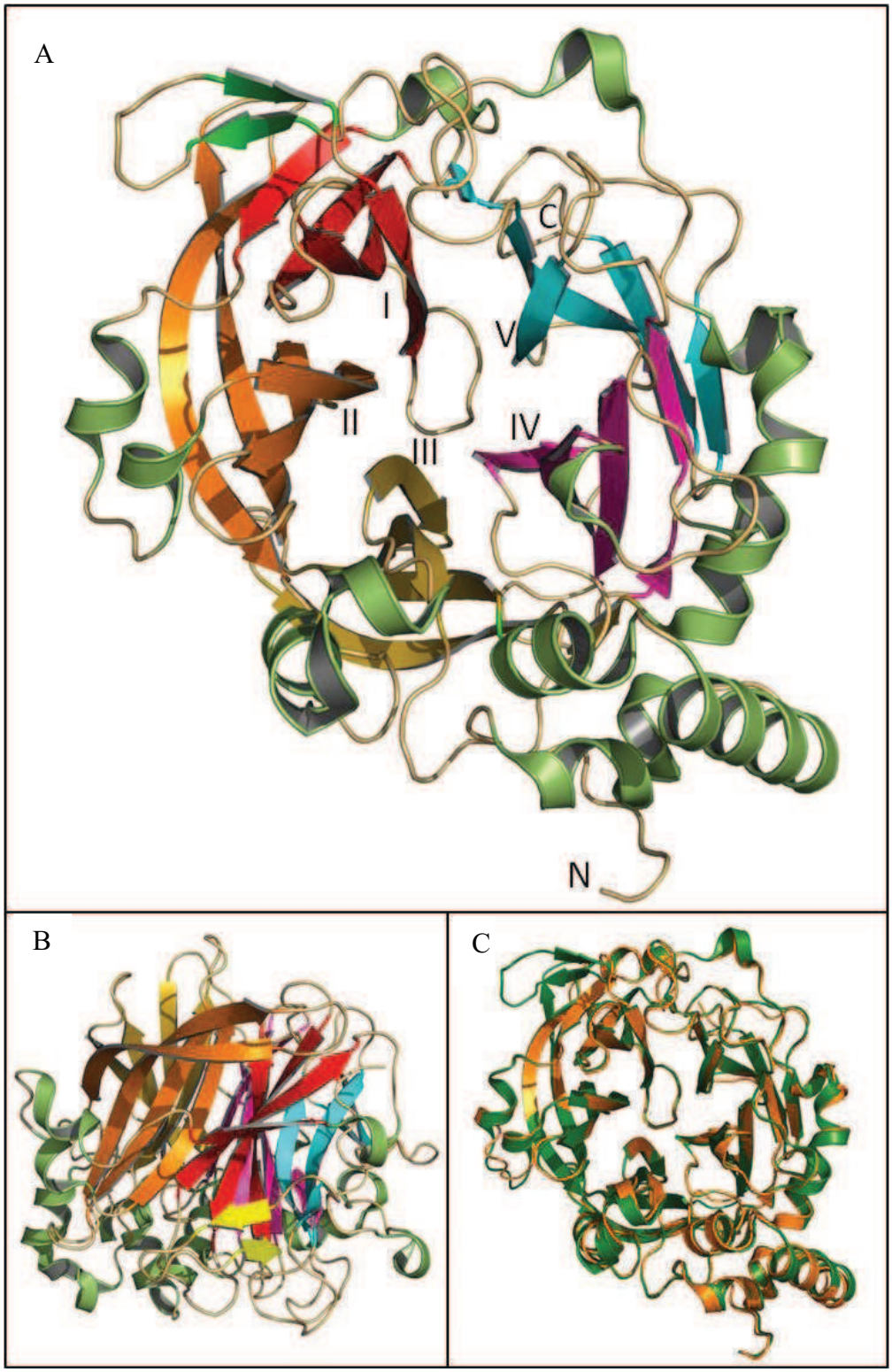

Figure 4-8: Ribbon representation of the overall structure of SacB variant D257A and superimposition with *B. subtilis* levansucrase.
(A, B) Two orthogonal views of the β-propeller fold of SacB from *B. megaterium*. The β-strands are colored according to their sequence succession (A) the view is along the active site tunnel. (C) Secondary structure-based superimposition of levansucrase from *B. megaterium* (green) and *B. subtilis* (orange), generated, using DALI (Holm and Park, 2000).

4.A.3.8 Active site of SacB

The active site is located at the bottom of the central cavity of the propeller. A funnel-like opening provides access to the deep negatively charged pocket towards the molecular surface. This pocket contains the three strictly conserved residues Asp95, Asp257 and Glu352 (Figure: 4-9) which constitute the catalytic triad. Glu352 acts as the general acid/base catalyst and protonates the glycosidic bond of sucrose while Asp95 acts as the nucleophil and forms an enzyme-fructosyl intermediate. Asp257 coordinates the fructofuranosyl moiety in position 3-OH and 4-OH. This catalytic triad in the active site of levansucrases (family GH 68) was described previously and proven by several mutagenesis studies (Chambert and Petit-Glatron 1991; Meng and Fütterer 2003; Martinez-Fleites *et al*., 2005; Homann *et al*., 2007; Meng and Fütterer 2008). Dividing the active site into subsites according to the terminology for active sites of glycoside hydrolases by Davies *et al*. (1997), levansucrases bind their substrate sucrose in subsite -1 (fructosyl residue) and +1 (glucosyl residue). Cleavage of sucrose takes place between subsites -1 and +1 (van Hijum *et al*., 2006) (Figure 4-9).

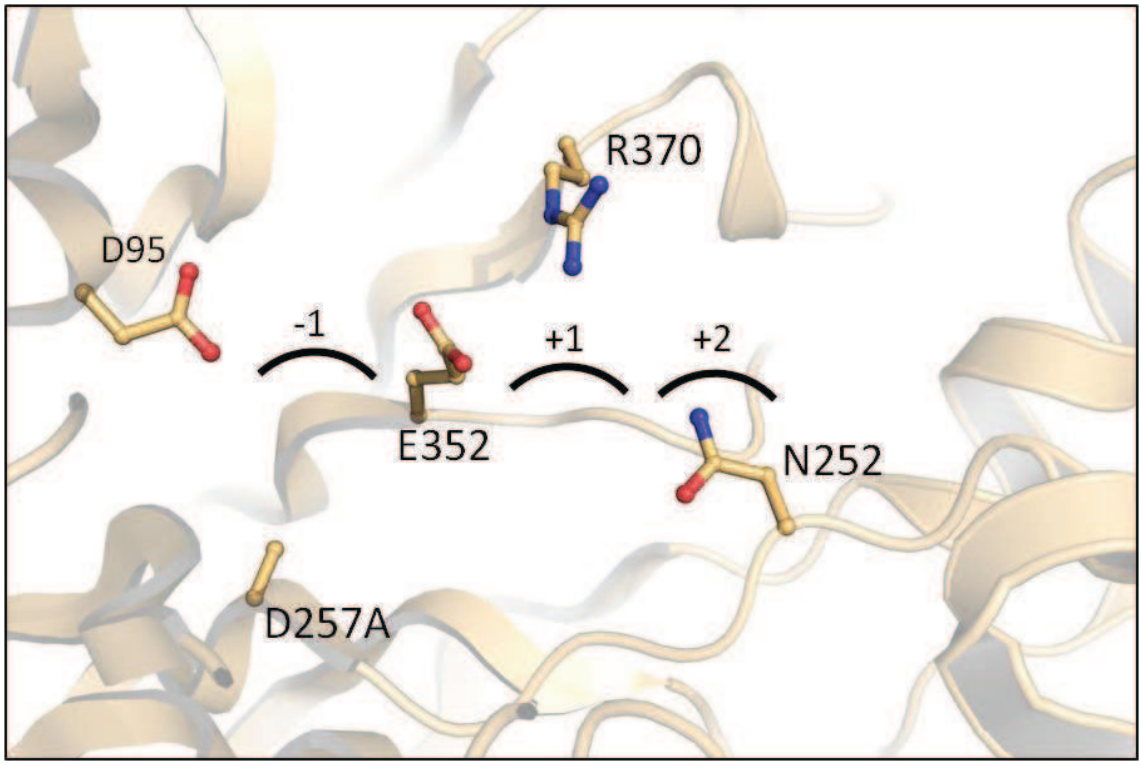

Figure 4-9: Close-up view of the active site of SacB D257A. Carbon, oxygen and nitrogen atoms are coloured light orange, red and blue, respectively. Proposed subsites for substrate binding are designated with arcs and numbers.

Arginine in position 370 is analogous to the previously described Arg360 of SacB from *B. subtilis* influencing the polysaccharide synthesis in subsite +1 (Homann, 2009).

4.A.4 Structural analysis of SacB variants with impact on the transfructosylation process

In previous investigations Homann and co-workers (2007) confirmed the crucial role of Asn252 in fructosyl transfer *versus* hydrolysis reactions. SacB variant N252A does not synthesize any polysaccharides while maintaining a high hydrolysis activity in the wild-type SacB range. It is found that Asn252 is not located in the active site. The question arises if the residue is directly involved in interactions with the fructan chain or if indirect effects, perhaps towards residues located in the catalytic site, are responsible for the abrogation of polymer formation. However, the structure of SacB variant N252A shows an intact active site compared with D257A (Figure 4-10). None of the residues of the catalytic triad changed its conformation.

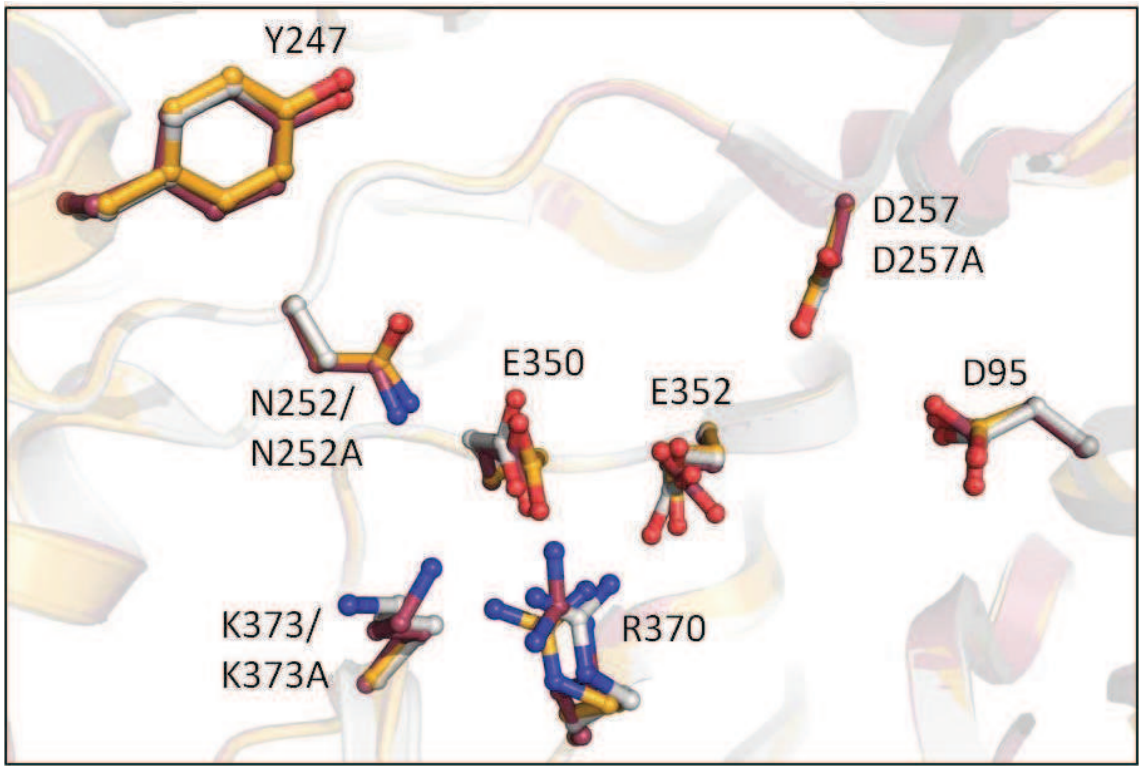

Figure 4-10: Superposition of SacB variants N252A (grey), K373A (light orange) and D257A (dark red). The structures of the indicate SacB variants show a maintained active site architecture.

Nevertheless, the orientation of residues Lys373 and Arg370 differs slightly in N252A. In variant N252A, residue Lys373 points in the direction of the alanine mutation in position 252. Exchanging Tyr247 to alanine and tryptophan, respectively shall give insights into potential protein-carbohydrate π-π stacking mechanisms in this location. Lys373 was exchanged to alanine. To clarify the role of these residues in the polysaccharide synthesis mechanism, the structures of SacB variants Y247A, Y247W and K373A were solved.

The superimposition of SacB variants Y247W, Y247A, K373A with D257A shows the conformationally unchanged active site and amino acids on the enzyme`s surface taking part in the transfructosylation process (Figure 4-11). However, the orientation of Arg370 is slightly different in the structure of the variant K373A.

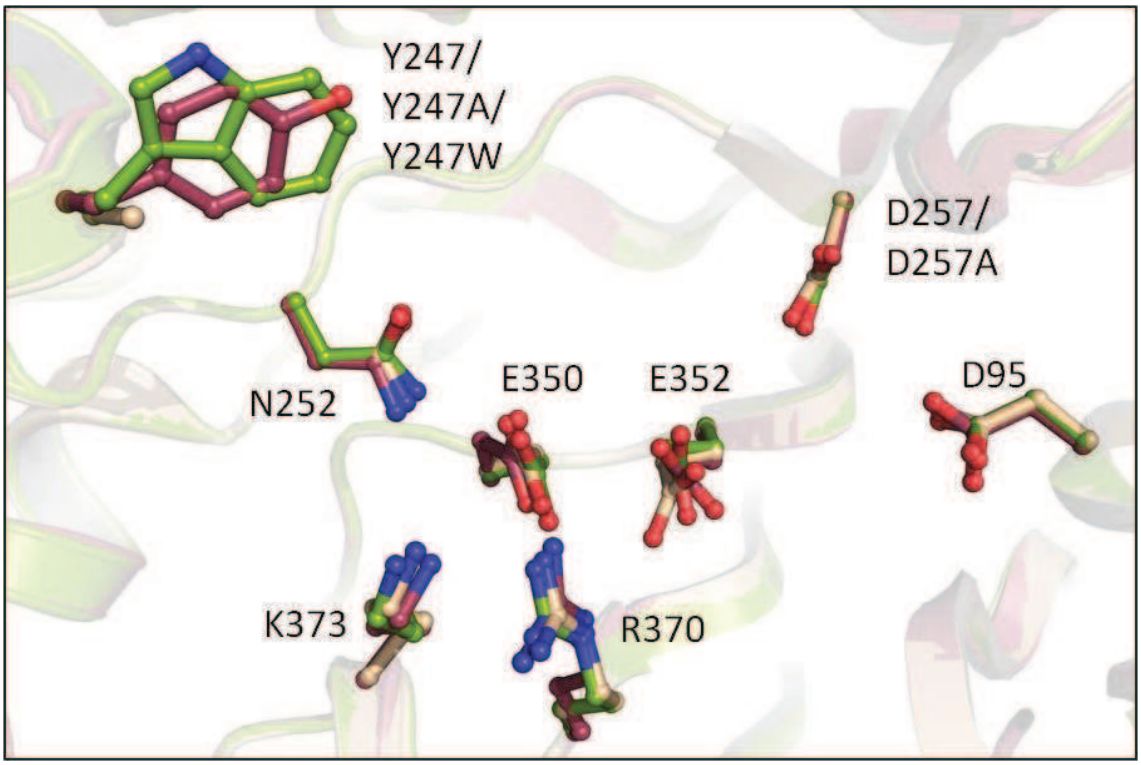

Figure 4-11: Superposition SacB variants Y247A (wheat), Y247W (green) and D257A (dark red). The structures of the indicated SacB variants show a maintained active site architecture.

4.A.5 Crystallographic analysis of potential SacB complexes with different ligands

All in all, during this thesis 42 complete data sets of the five SacB variants were collected in order to obtain ligand-bound structures. For each variant only one ligand-free structure was completely refined. As a possible aim of this thesis was the analysis of a complex

between substrate, product or acceptor substrates and SacB by X-ray analysis, all collected data sets were processed, but structure refinement only proceeded to that point where residual electron density would allow such a potential ligand to be identified. However, in no case such a complex could unambiguous be identified.

4.A.6 Calcium binding site of SacB

Analysis of the crystal structure of *B. subtilis* levansucrase has provided evidence for the presence of a metal ion, most likely Ca^{2+} (Meng and Fütterer, 2003). Sequence alignments between family GH68 members revealed that the residues involved in binding of the calcium are strictly conserved in most enzymes of gram-positive bacteria like *B. megaterium*, but are absent in proteins of Gram-positive bacteria (Ozimek *et al*., 2006; van Hijum *et al*., 2006).

In *B. megaterium* SacB the metal binding site is located between helices α3 and α5. The binding site contains a Ca^{2+} ion and has a pentagonal bipyramidal geometry (*B*-factors between 4.11 and 23.58 $Å^2$) (Figure 4-12).

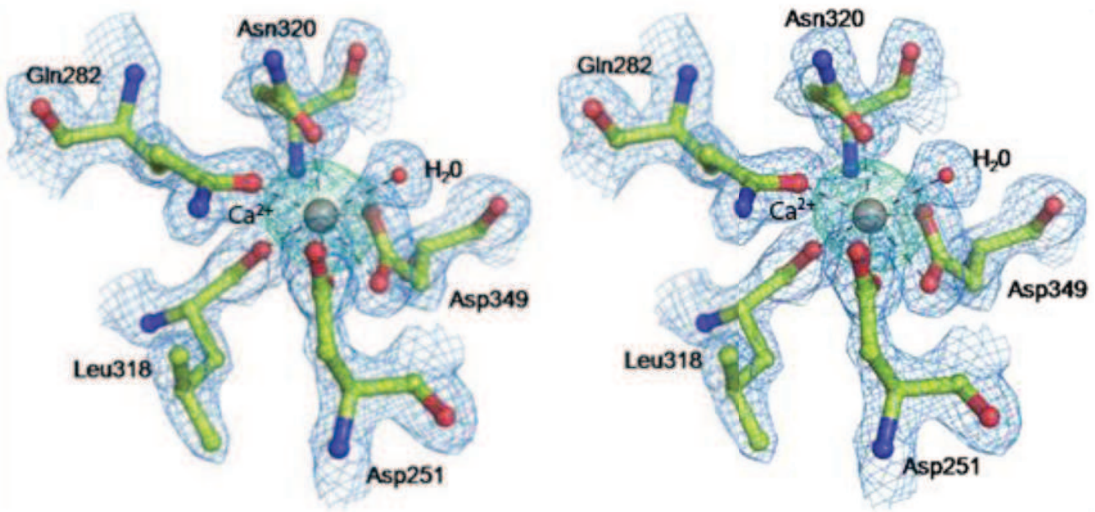

Figure 4-12: Electron density of the Ca^{2+} binding site of *B. megaterium* levansucrase (stereo view). The σ_A-weighted $2F_o$-F_c density (blue mesh) is contoured at 1.5 σ. Grey sphere depicts Ca^{2+}. Selected residues are shown as sticks. F_o-F_c difference map (green mesh, 3.0 σ) was calculated by omission of Ca^{2+}.

Refinement was carried out with an occupancy of 1.0. The Ca^{2+} ion is coordinated by five oxygen ligands, contributed by SacB, with distances between 2.3 and 2.6 Å: Asp251 (2.3 Å), Gln282 (2.3 Å), Leu318 (2.4 Å), Asn320 (2.4 Å) and Asp349 (2.5 and 2.6 Å). In

addition to these five residues, the central calcium ion forms a coordinate bond with a water molecule (2.5 Å). In total the metal has 7 ligands (carboxylate group of Asp349 is bidentate).

4.A.7 Binding of PEG to the active site pocket

In the final stages of refinement, inspection of the difference map (F_o-F_c) showed clearly elongated density at the entrance of the active site pocket, for all variants except for D257A. First attempts to fit this density with one of the soaked compounds like 1-kestose or 6-kestose failed. As this electron density could also be observed for data of crystals which were not treated with soaking compounds, the electron density was interpreted as a polyethylene glycol (PEG) molecule, which was derived from the crystallization solution. Figure 4-13 shows a bound PEG molecule in the active site cavity of variant N252A.

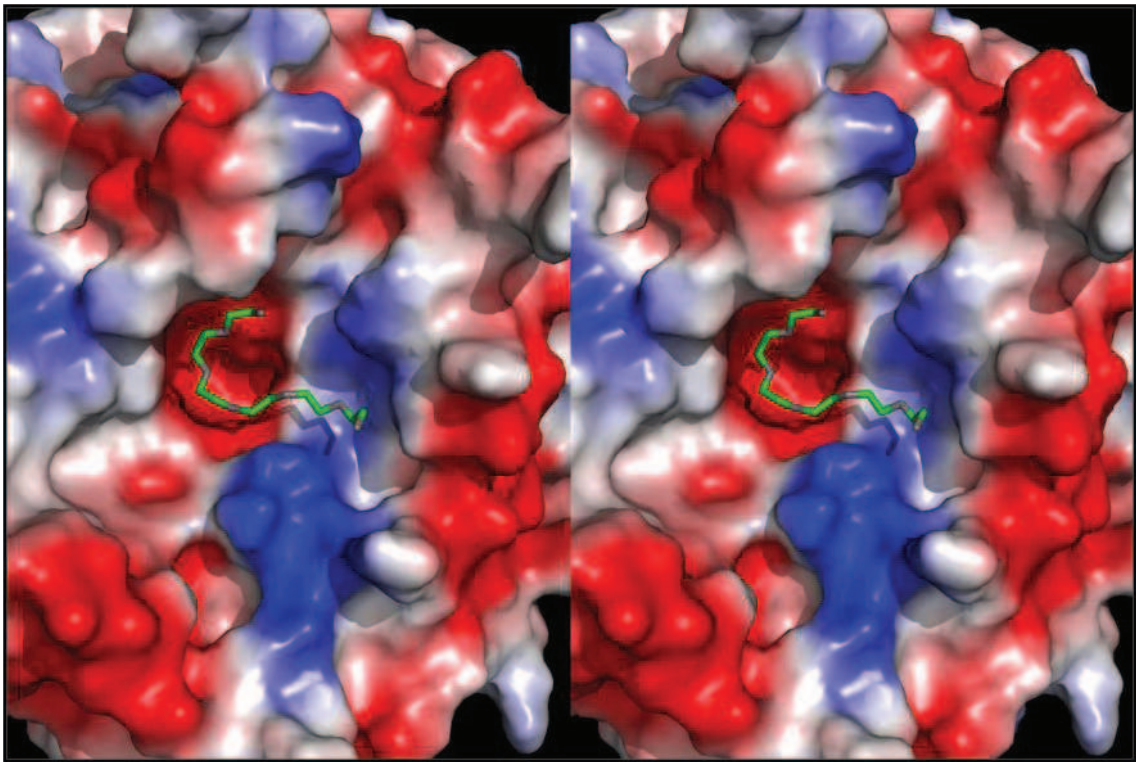

Figure 4-13: Stereo view SacB N252A (chain B) with bound PEG molecule in the active site cavity. The surface presentation shows the charges of the residues involved (red = negative; blue = positive; white = uncharged). The bound PEG molecule is shown in green (carbon) and grey (oxygen).

Several ethylene glycol polymers with a different number of non-hydrogen atoms could be modeled into the active site pockets of the variants. The number of non-hydrogen

atoms of these PEG fragments varied between seven and 19 atoms, which corresponds to PEG molecules of between approximately 100 and 300 Da. It is well known that commercially available PEG reagents contain a mixture of molecules of different molecular masses, which would explain the electron density for a lower PEG mass than used for crystallization. Another possibility is that only parts of the molecules are ordered in the crystal and thus visible in the electron density map (Oliveira *et al*., 2006).

Almost identical electron density was found in almost all variants. For all four molecules in the asymmetric unit of variant N252A (chain A-D) in total six PEG molecules or fragments could be identified in the active site pocket. Figure 4-14 shows a PEG molecule bound in the active site cavity of variant N252A (chain B). Seven PEG molecules were found in variants K373A and Y247W and one molecule in Y247A.

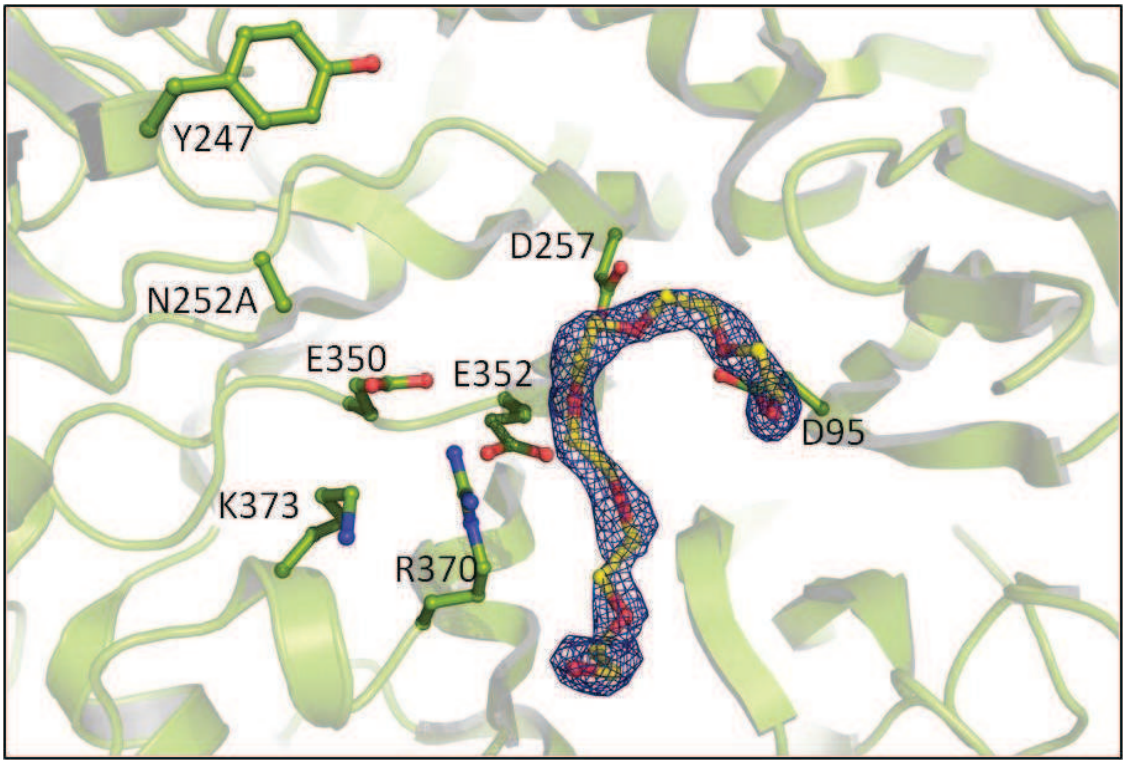

Figure 4-14: PEG molecule bound in the active site pocket of SacB N252A, chain B. The σ_A-weighted $2F_o$-F_c density (blue mesh) is contoured at 1.0 σ.

The active site pocket perfectly accommodates a PEG molecule that lies between Trp94 and Pro414 and is involved in several hydrophobic interactions. The ligplot diagram (Wallace *et al*., 1995) is given in Figure 4-15, highlighting hydrophobic contacts between the PEG molecule and residues of variant N252A. Residues of the catalytic triad are not

directly involved in interactions, but surrounding residues Trp94, Leu118, Trp172 and Pro414.

The active site of D257A only showed weak density for either water or perhaps several short PEG molecules. A superposition revealed that the electron density in D257A was partly consistent with the density of the other variants in this region. However, the PEG molecules did not fit very well, thus water molecules were introduced into the residual density of variant D257A.

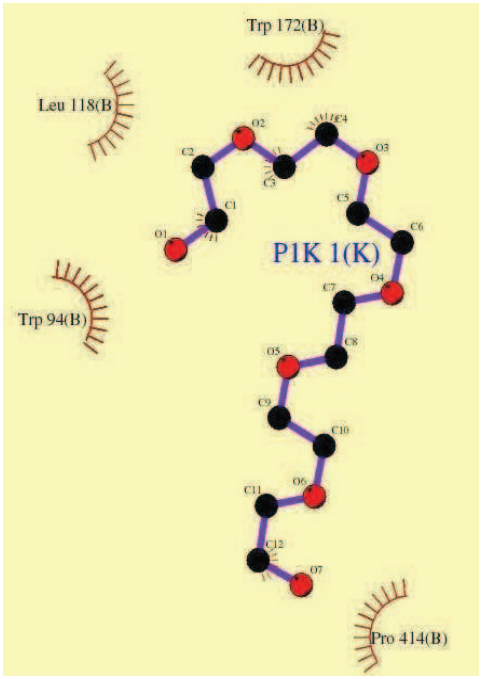

Figure 4-15: Ligplot diagram of SacB N252A showing hydrophobic interactions between the enzyme and a bound PEG molecule. Hydrophobic contacts are represented by an arc with spokes radiating towards the ligand atom they contact.

4.A.8 Validation of the SacB models

Final refinement statistics for the five SacB variant models are given in Table 4-4. The refinement progress was continuously judged by monitoring the values of R_{cryst} and R_{free} throughout model building. The five structures were refined to R_{cryst}/R_{free}- values of 16.1/21.8 (D257A), 18.1/20.8 (N252A), 20.2/21.3 (K373A), 22.2/23.0 (Y247A), and 20.1/21.6 (Y247W), respectively. Stereochemical parameters such as r.m.s. deviations of bond lengths and angles were in a normal range and showed no significant deviations

from standard structures. MOLPROBITY web server (Davis *et al*., 2007) was used for all-atom contact analysis, confirming no serious clashes. Torsion angles of the protein backbones were plotted in a Ramachandran diagram calculated with the program RAMPAGE (Lovell *et al*., 2003) (Figure 4-17). Almost all residues of the final models were grouped in the favored or the allowed region of the Ramachandran plot, with the exception of Thr441 in variant D257A, Ser125 and Asn87 in variant N252A and again Ser125 in variant Y247W. Manual inspection of these outliers revealed that residue Thr441 is well defined in the electron density (Figure 4-16), whereas the other remaining outliers are located in less well-defined loop regions. However, Thr442 is not located near the active site region and therefore has probably no influence on the enzyme´s activity.

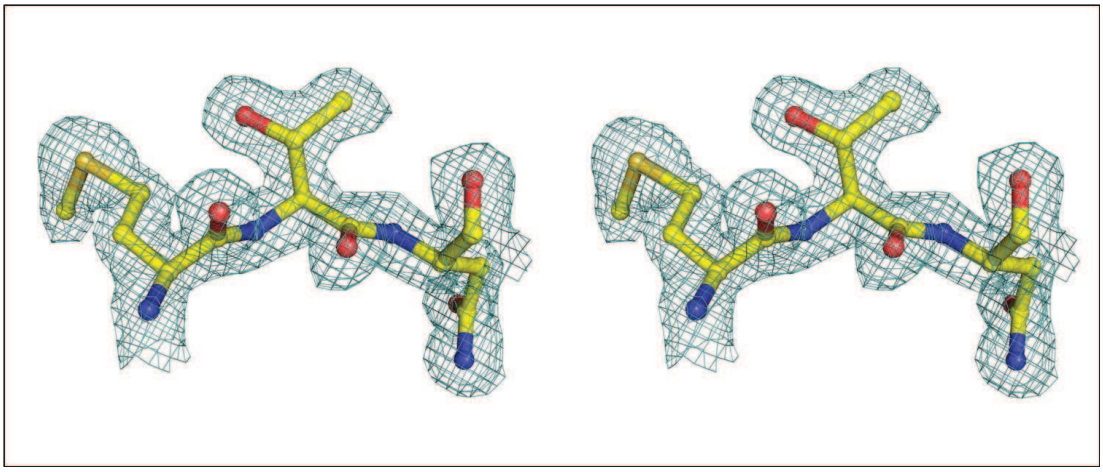

Figure 4-16: Stereo view of Ramachandran outlier Thr441 in SacB D257A. The residue reveals an excellent fit to the refined $2F_o$-F_c map (blue mesh, 1.0 σ).

The first five N-terminal residues and the four last C-terminal residues were not included in the model due to the lack of well-defined electron density. These residues are presumably flexible and disordered.

Table 4-4: Structure refinement statistics.

Refinement data	D257A	N252A	K373A	Y247A	Y247W
Spacegroup	$P2_12_12_1$	$P2_1$	$P2_1$	$P2_1$	$P2_1$
Resolution (Å)	24.7-1.9	47.6-2.0	47.8-1.8	47.8-2.0	47.7-1.9
Molecules/AU	1	4	4	4	4
R_{cryst}	16.1	18.1	20.2	22.2	20.1
R_{free}[a]	21.78	20.8	21.3	23.0	21.6
Number of atoms					
protein	3601	14182	14260	14184	14235
PEG molecules	-	6	7	1	6
water molecules	657	1156	1457	1277	1094
calcium ions	1	1	1	1	1
magnesium ions	2	-	-	-	-
citrate molecules	2	-	-	-	-
sulfate molecules	-	6	7	1	6
Average *B*-factor(Å^2)					
Protein	7.1	25.0	12.3	13.4	16.0
PEG molecules	-	41.4	24.8	20.9	34.9
water molecules	20.8	34.4	21.3	24.7	23.9
calcium ions	4.1	23.6	7.9	13.0	14.0
magnesium ions	17.3	-	-	-	-
citrate molecules	14.5	-	-	-	-
sulfate molecules	-	40.2	31.1	31.9	38.0
r.m.s.d. from ideal[b]					
bond length (Å)	0.015	0.007	0.006	0.006	0.006
bond angle (°)	1.316	1.028	1.023	0.897	0.992
Ramachandran statistics [number of residues/%]					
in favored region	434/97.3	1722/96.5	1729/96.9	1722/96.5	1724/96.6
in allowed region	12/2.5	60/ 3.4	55/3.1	61/3.4	60/3.4
in outlier region	1/0.2	2/0.1	0	1/0.1	0
PDB code	3om2	3om5	3om4	3om6	3om7

[a] R_{free} calculated using 5 % of total reflections omitted from refinement

[b] r.m.s.d., root mean square deviation

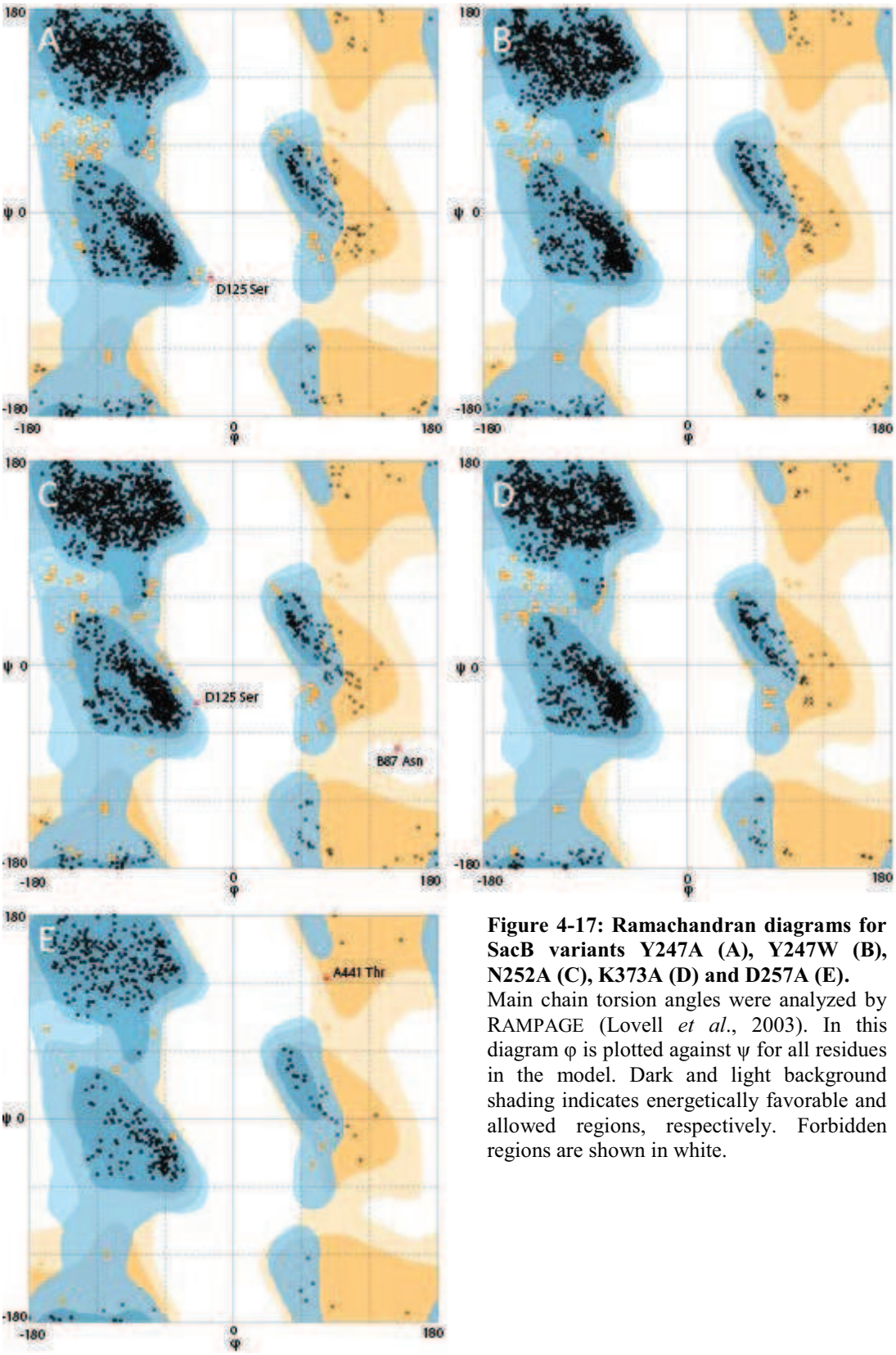

Figure 4-17: Ramachandran diagrams for SacB variants Y247A (A), Y247W (B), N252A (C), K373A (D) and D257A (E).
Main chain torsion angles were analyzed by RAMPAGE (Lovell *et al.*, 2003). In this diagram φ is plotted against ψ for all residues in the model. Dark and light background shading indicates energetically favorable and allowed regions, respectively. Forbidden regions are shown in white.

4.A.9 Crystallization in complex with substrates, products and inhibitors

Once well diffracting native crystals of enzymes are obtained, the addition of substrates, products or potential inhibitors is needed to obtain a molecular complex of the target protein with the ligand. Particular for the investigation and understanding of the enzyme´s reaction mechanism complexes are required. In order to get these protein-ligand complexes, several co-crystallization- and soaking experiments with different ligands have been applied.

Co-crystallization experiments

Co-crystallization experiments of wildtype SacB and its variants N252A, K373A, D257A, Y247W, Y247A and E352A were performed with the fructosyltransferase products 1-kestose, 6-kestose and with a potential inhibitor of SacB allosucrose (α-D-allopyranosyl β-D-fructofuranoside). In addition to setups with the products and the inhibitor, the inactive SacB variants E352A and D257A were also incubated with the substrates of levansucrase sucrose and raffinose.

Allosucrose, 1-kestose, and 6-kestose were kindly provided by the group of Prof. Dr. Jürgen Seibel (University of Würzburg). The final concentration of the particular ligand was chosen to be in a 2 to 10 fold molar excess to the protein to achieve reasonably binding site occupancy. However, crystal growth could not be observed under the conditions using protein concentrations similar to the non co-crystallization experiments. In fact, almost all of the prepared crystallization drops neither showed any signs of crystallization nor precipitation of the protein. Repeating these experiments with higher protein concentrations (10, 11, 12 and 20 mg/mL) resulted in formation of heavy precipitate after a few days (not shown).

Crystal soaking experiments

Soaking experiments were performed with crystals of the variants N252A, K373A, D257A, Y247A, and Y247W by the addition of allosucrose, 1-kestose, 6-kestose raffinose and sucrose.

Crystals of each variant were soaked for 1, 2, 5, 10, 30 and 60 min, respectively, in mother liquor (20 % (w/v) PEG 1000, 0.2 M lithiumsulfate, 0.1 M Na-phosphate pH 4.1 with and without 1 mM $CaCl_2$) containing 1, 10, 20, 50, 100 and 150 mM of the soaking compound, respectively. For cryoprotection the mother liquor was mixed with glycerol to a final concentration of 25 % and supplemented with the soaking compound in the respective concentration. As a consequence, treatment with these substances always caused a substantial decrease in crystal quality. The higher the concentration and the longer the soaking time, incubated crystals gradually dissolved and disappeared after a while. X-ray data of short-soaked crystals were collected at beamlines at DESY, BESSY and at the ESRF. However, either crystal quality was affected too strongly by soaking to allow for successful processing of collected data sets (Figure 4-18) or only ambiguous electron density corresponding to one of the soaked substances was found in the structures of the crystallized variants.

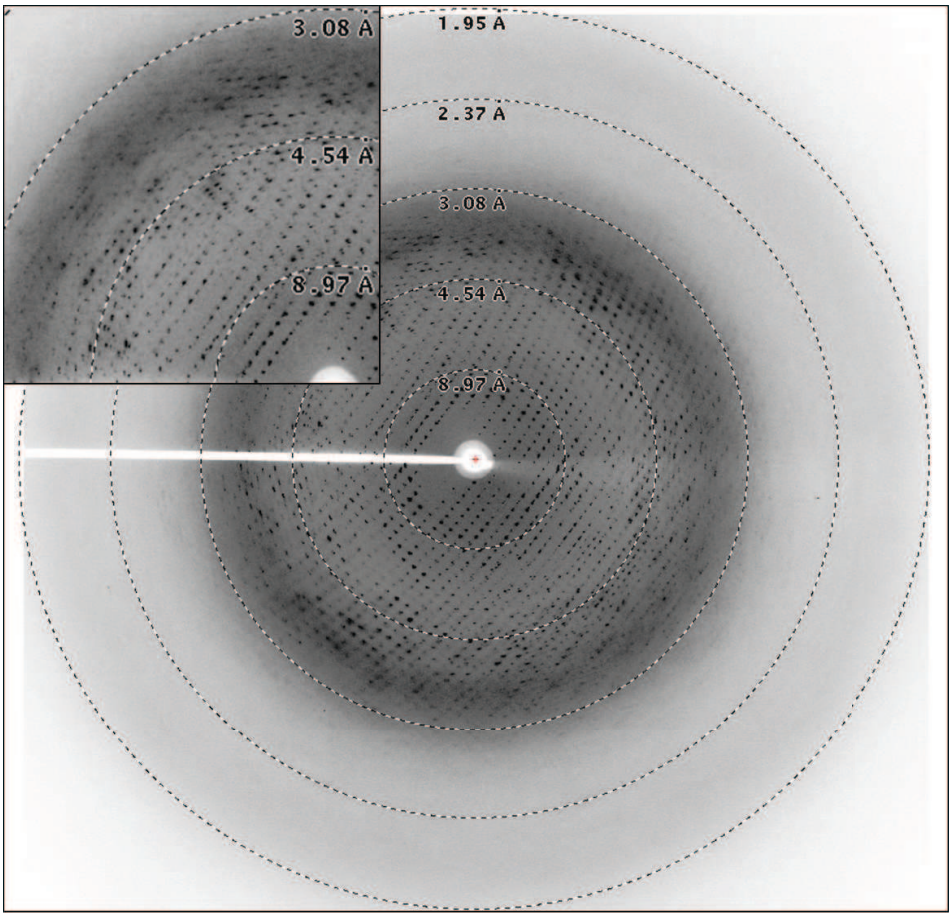

Figure 4-18: Diffraction image of a SacB Y247W. Crystal after treatment with a solution of 50 mM allosucrose for 10 min. The diffraction image shows signs of crystal damage after soaking experiments. Upper left inset shows a close-up view of the diffraction image.

In order to protect the crystals and keep them intact, crystals were also transferred to a series of solutions containing increasing concentrations of sugars. Crystals of SacB variant N252A were transferred first to mother liquor supplemented with 25 mM 1-kestose (4 min), then soaked for 8 min in mother liquor containing 50 mM 1-kestose followed by 10 min of soaking in mother liquor containing 100 mM. However, the results were the same as for the crystals which were transferred at once to soaking solutions – they deteriorated and dissolved after a few minutes.
Despite extensive efforts, diffraction quality crystals in complex with either products or sucrose analogue could not be obtained by soaking experiments.

Protein purification in the presence of ligands

In another effort to prepare SacB and inactive variants in complex with ligands for X-ray diffraction studies, sucrose and allosucrose were either included in the chromatography buffers (10 mM to 100 mM) or ligands, such as 6-kestose, 1-kestose and allosucrose (10 mM to 50 mM) were directly given to the protein samples (for the purification of variants N252A, K373A, Y247A, Y247W) before loaded onto the chromatography columns. Purification in the presence of ligands was successful and subsequent crystallization trials yielded crystals similar to those, not purified with buffers mixed with any ligands. These crystals diffracted up to resolutions ranging from 2.3 to 3.0 Å. However, structure determination revealed that no complex crystals were obtained. Analysis by mass spectrometry of the protein samples purified in the presence of ligands could not confirm a complex as well.

4.A.10 Crystal transfer experiments

In order to obtain detailed information of the interactions of SacB variants with several ligands, a protein-ligand complex is indispensable. During structure refinement of SacB variants N252A, K373A, Y274A, and Y247W, the Fourier difference map clearly showed elongated electron density in the active site cavity, into which PEG fragments were modeled (see section 4.A.7). However, the bound PEG fragments could prevent the binding of substrates and products to the active site cavity. Therefore, a transfer of SacB

crystals from the mother liquor containing PEG 1000 to a solution with another precipitant was performed, in order to remove the PEG fragments from the binding site of the protein and to keep the crystal intact.

According to the protocol described by Schreuder and co-workers (1988), a series of solutions containing increasing concentrations of alternative precipitants were pipetted into the reservoirs of a hanging-drop setup. In order to keep the crystals intact and to avoid the attraction of water the chemical potential or the partial molar Gibbs free energy of the water of two different mother liquors should be the same. For determining the chemical potential in the different solutions, vapor equilibration experiments were performed with solutions containing following alternative precipitants: NaCl (0.5 – 3.5 M), $AmSO_4$ (0.2 – 2.2 M) and MPD (10 – 60 %). From the results of the equilibration studies it could be concluded that a NaCl solution of 2.0 to 2.5 M and a solution of 1.5 M $AmSO_4$ are approximately equivalent to the crystallization condition, comprising 20 % PEG 1000, 0.2 M lithiumsulfate and 0.1 M Na-phosphate-citrate pH 4.1. When SacB crystals were transferred at once to the NaCl solution, they deteriorated. Gradually replacement of PEG 1000 by NaCl did not result in visually damage of the crystals, but the transferred crystals did not diffract as well as the origin crystals (7.0 – 10.0 Å) using synchrotron radiation.

4.A.11 Alternative crystallization method to avoid PEG binding

In order to obtain SacB crystals without PEG molecules bound in the active site of the enzyme, crystallization of SacB according to the method of Meng and Fütterer (2003) was applied. Meng and Fütterer crystallized the homologue levansucrase SacB from *B. subtilis* by a first dialysis step against deionized water to get microcrystalline precipitate. In a second step sizeable crystals were obtained by microdialysis of the protein (60 mg/mL) against 1 mM diammonium phosphate and 1 mM sodium acetate. SacB from *B. megaterium* could be concentrated to ~60 mg/mL. However, even by testing protein concentrations between 20 to 70 mg/mL neither crystals nor microcrystals were obtained. The protein precipitated after two days in a first dialysis experiment. Using this precipitate for a second microdialysis step also did not result in the formation of crystals.

4.A.12 Thermal shift assay

The thermofluor shift assay is a high-through put method to detect potential interactions between ligands and proteins by measuring thermal stability of the protein. The aim was to investigate factors affecting the stability of SacB indicating promising crystallization conditions for the generation of complex structures. In order to determine the optimum condition for the assay, three different concentrations of SacB WT and variant K373A (0.5, 1.0 and 2.0 mg/mL) and three concentrations of the fluorescent reagent Sypro® Orange (100x, 50x and 10x) were tested before conducting the thermal shift assay. For the subsequent assay, 1.0 mg/mL of SacB WT and variant K373A and 100 x of Sypro® Orange were used. Stability changes of the wildtype SacB were measured in the presence of the putative inhibitor allosucrose. 1-kestose was used for the assay wit SacB K373A. The protein was pipetted into a 96-well plate and heated from 20°C to 95°C in 0.5°C steps. The fluorescent signal was recorded. In the absence of allosucrose and 1-kestose, respectively, the melting temperature for both samples was at 39.5°C (Figure 4-19 and 4-20). For wildtype SacB a relatively small increase of the melting temperature to 43.0°C was observed in the presence of 50 mM allosucrose. Measurements of the melting temperature in the presence of allosucrose concentrations above 50 mM showed a subsequent decrease of the melting temperature down to 42.0°C (100 mM) and 38.0°C (150 mM) compared to 43°C in the presence of 50 mM allosucrose (Figure 4-19).

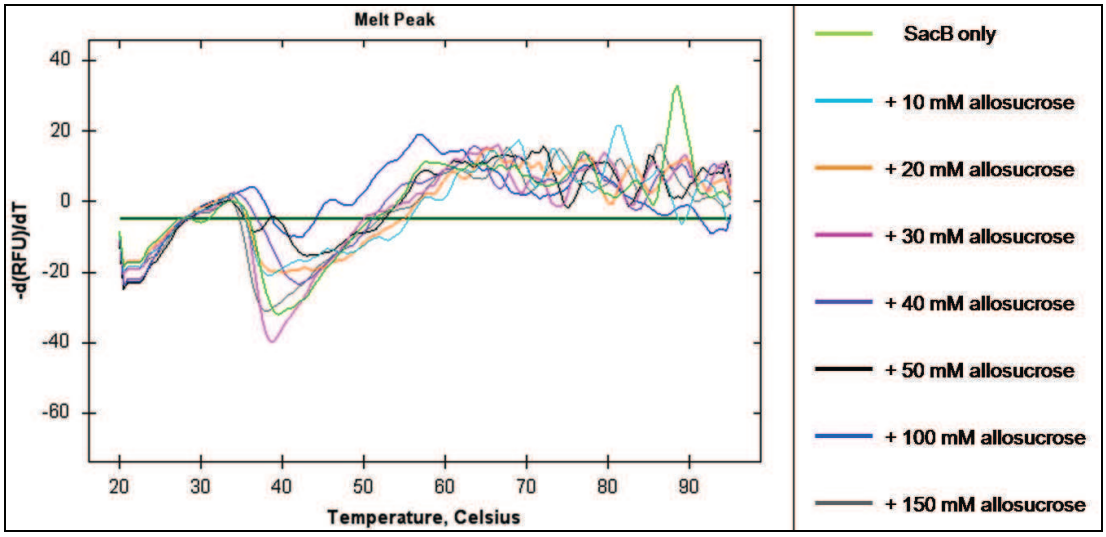

Figure 4-19: Thermofluor analysis of interaction between wt-SacB with allosucrose.

In the presence of 1-kestose the melting temperature of variant K373A was increased by 3.5 °C to 43.0°C. Higher concentrations of 1-kestose (100 mM and 150 mM) led to a decrease of the T_m to 40.0 °C and 38.5 °C, respectively (Figure 4-20). The measurements were always repeated two times in order to prove significance of the results.

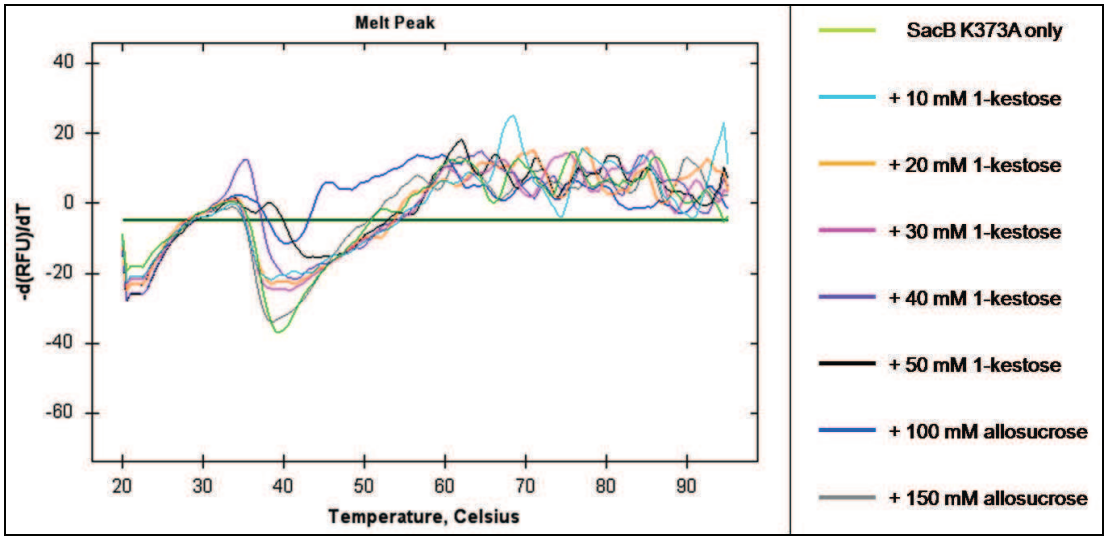

Figure 4-20: Thermofluor analysis of interaction between K373A with 1-kestose.

The results of the thermofluor shift assays indicated a stabilization of the protein samples in the presence of 1-kestose and allosucrose in concentrations of about 50 mM. However, soaking experiments as well as co-crystallization experiments using concentrations of 1-kestose and allosucrose, respectively, in the range of 10 to 150 mM still did not lead to SacB-structures in complex with detectable amounts of ligands.

4.B Glycosyltransferase R from *Streptococcus oralis*

The second part of the thesis involves the investigation of the glucan binding domain (GBD) of the dextransucrase GtfR from *Streptococcus oralis*. Some of the experiments presented below were performed as part of the diploma thesis of Dipl. Biol. Maria Ebbes (generation of *gbd*-constructs, purification of the GBD and first crystallization trials).

4.B.1 Identification of the Glucan Binding Domain of GtfR

Several studies showed that the C-terminal domain of bacterial glucansucrases is responsible for the binding of glucan. Therefore, it is also named glucan binding domain (GBD) (Kato and Kuramitsu, 1990). This domain is characterized by a specific pattern of repeated tyrosine (Y) and glycine (G) residues in their primary structure, named YG repeats. To identify the boundary of the GBD of GtfR from *S. oralis* a BLAST search was performed. BLAST (http://www.expasy.org/tools/blast/) is a bioinformatic tool that enables the comparison of a query sequence within a vast database, aiming at the identification of library sequences that resemble the query sequence above a given threshold. A search for homologous sequences GtfR was performed selecting the full length GtfR as the query. Figure 4-21 shows an excerpt of the BLAST search output.

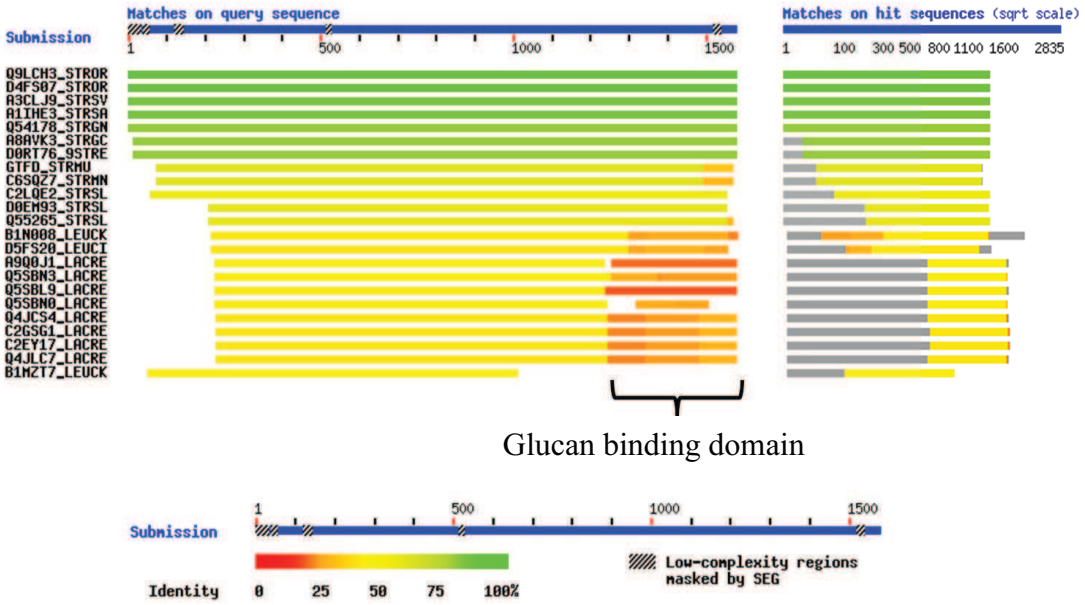

Figure 4-21: BLAST search performed with GtfR sequence.

As expected, BLAST indentified several bacterial glucansucrases sharing similar sequences and putative glucan binding domains. In order to have a more precise estimate of the C-terminal GBD and its YG repeats, a sequence analysis of the C-terminus using RADAR was performed. RADAR (www.ebi.ac.uk/Tools/Radar; Heger and Holm, 2000) is a web server for segmenting an amino acid sequence into repeats. In Figure 4-22, the additional manually checked 13 YG motifs are shown.

```
1169   G N W Y Y F D K R G Y L V T G A H E I D
1189   G K H V Y F L K N G I Q L R D S I R E D E N
1210   G N Q Y Y Y D Q T G A Q V L N R Y Y T T D G
1232   Q N W R Y F D A K G V M A R G L V K I G
1252   D G Q Q F F D E N G Y Q V K G K I V S A K D G
1275   K L R Y F D K D S G N A V I N R F A Q G D N P
1298   S D W Y Y F G V E G A K L T G L Q K I G
1318   Q Q T L Y F D Q D G K Q V K G K I V T L S D K
1341   S I R Y F D A N S G E M A V G K F A E G A K
1363   N E W Y Y F D K T G K A V T G L Q K I G
1383   K Q T L Y F D Q D G K Q V K G K V V T L A D K
1406   S I R Y F D A D S G E M A V G K F A E G A K
1429   N E W Y Y F D Q T G K A V T G L Q K I D
1449   K Q T L Y F D Q D G K Q V K G K I V T L S D K
1472   S I R Y F D A N S G E M A T N K F V E G S Q
1494   N E W Y Y F D Q A G K A V T G L Q Q V G
1514   Q Q T L Y F T Q D G K Q V K G K V V D V N G
1536   V S R Y F D A N S G D M A R S K W I Q L E D
1558   G S W M Y F D R D G R G Q N F G R N
```

Figure 4-22: Alignment of the "YG" sequences from the GBD of GtfR from *S. oralis*. The identifier for these sequences was the occurrence of glycine four to six residues after a cluster of one or more aromatic residues.

Once the approximate starting point of the GBD was defined, the position of the truncation was chosen so as to avoid cutting onto the first YG motif. Secondary structure prediction was performed to identify an interdomain region between the GBD and the

catalytic domain. Therefore, the secondary structure prediction tool, Phyre (Protein Homology/analogy Recognition Engine; http://www.sbg.bio.ic.ac.uk/phyre/; Kelley and Sternberg, 2009) was used. Phyre combines different secondary structure prediction programs such as PSIPRED (Bryson *et al*., 2005), JNet (Cole *et al*., 2008), SSPRO (Cheng *et al*., 2005) and Consensus (Prasad *et al*., 2003). Figure 4-23 displays the structural prediction obtained and the predicted region chosen as the N-terminus of the construct (residues 1169 – 1575).

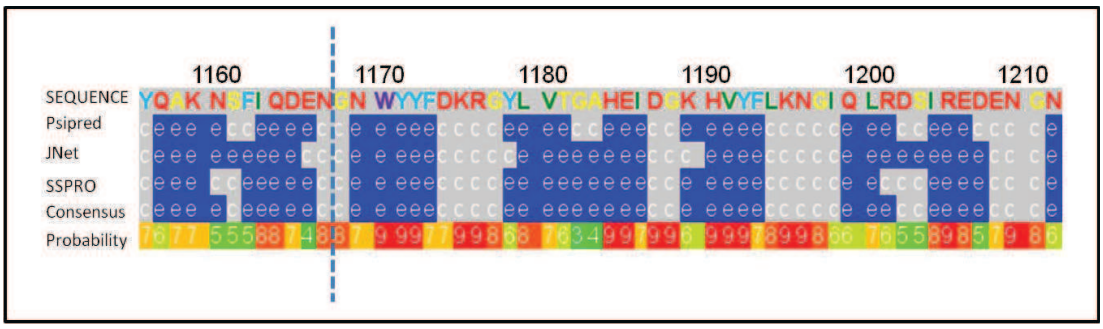

Figure 4-23: Structure predictions performed by Phyre. Dotted line indicates the chosen domain boundary. Indicated are predictions by Psipred, JNet, SSPRO and Consensus. c: coiled region; e: beta sheet.

4.B.2 Cloning strategy for GBD constructs

The gene of full length GtfR had previously been cloned into the vector pBlueSkript II KS by Dr. Bernd Hofer (HZI). This vector was used as template for the amplification of DNA sequences coding for constructs of the GBD. For expression of *gbd* in *E. coli*, several expression vectors were constructed, including pET28c, pETM30, pQE60 and pGEX-6P-1. These vectors provide the expressed protein of N- and C-terminal His- and GST-tags (Table 4-5). The genes were amplified by PCR, introducing restriction enzyme cleavage sites.

Generated constructs were controlled by test restrictions and partial sequencing. The *gbd*-insert was successfully cloned into all four vectors.

Table 4-5: Vectors used for the production of GBD-constructs

Vector	Tag	Protease cleavage site	Restriction sites in insert and recipient
pET28c	His_6-tag (N-terminus)	Thrombin	NdeI/BamHI
pETM30	GST-tag (N-terminus)	TEV	NcoI/BamHI
pQE60	GST-tag (N-terminus)	TEV	NcoI/BamHI
pGEX-6p-1	His_6-tag (C-terminus)	PreScission®	BamHI/NotI

4.B 3 Production of GBD in *E. coli*

Test expression of GBD constructs in *E. coli*

To determine an optimal expression procedure, small-scale test expression tests were carried out under different conditions. Samples were collected at different time points and soluble and insoluble fractions were monitored by SDS-PAGE. Figure 4-24 depicts the expression of pET28c-gbd in *E. coli* BL21 CodonPlus cells during an o.n. expression at 16°C and 25°C, respectively. Unlike the other three constructs, the expressed fusion protein with a MW of 48.2 kDa was clearly visible in an SDS-Gel. The best expression result of the construct was obtained o.n. at 25°C and after induction by 0.5 mM IPTG.

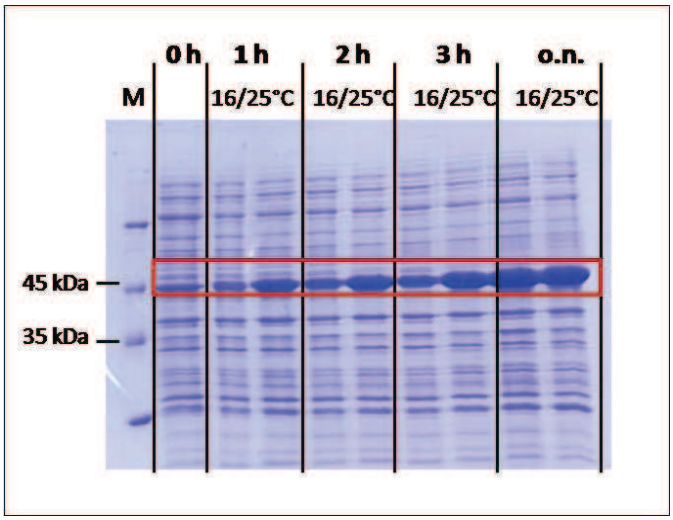

Figure 4-24: Test expression of His6-GBD in *E. coli* BL21 CodonPlus cells. Cell extract after 0, 2 and 3 h and overnight (o.n.) at 16°C and 25°C. Red box: bands at ~ 45 kDa containing His_6-GBD. M: Marker

Preparative expression and purification

His_6-GBD was purified from *E. coli* BL21 CodonPlus cells grown in 2 L LB-medium. Typically, about 7 mg per liter of bacterial culture were obtained of His_6-tagged GBD. A combination of several different chromatography techniques was chosen to achieve the purification of overexpressed gbd. Initial purification of the ~48 kDa His_6-GBD was performed by means of an affinity chromatography step. Coupling to a nickel (II)-nitrilotriaceticacid (Ni-NTA) resin was followed by several washing steps and elution with increasing imidazole concentrations. For later crystallization trials, the His_6-tag was cleaved by the addition of thrombin followed by re-chromatography with Ni-NTA to remove the His_6-tag. The final elution fractions already showed high degree of purity when analyzed by SDS-PAGE (Figure 4-25).

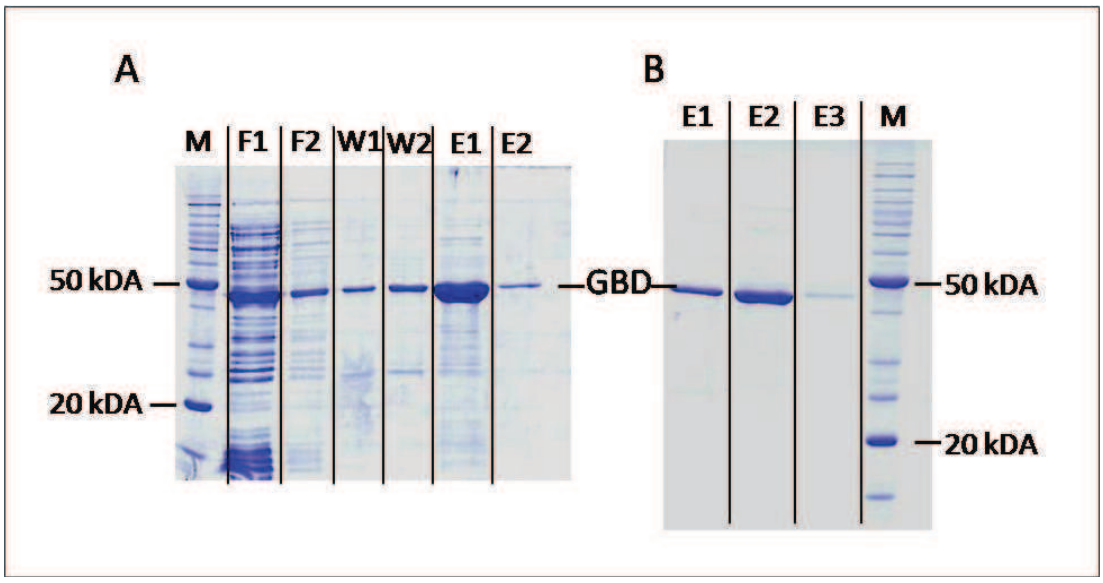

Figure 4-25: Purification of His_6-GBD via Ni-NTA affinity chromatography. Coomassie brilliant blue stained 12 % SDS-PAGE. (A) unbound (F1, F2), wash- (W1, W2) and elution fractions (E1, E2) from the affinity chromatography steps performed with *E. coli* cell extract containing His_6-GBD. The first lane contains the protein marker (M). (B) Elution fractions (E1-E3) of GBD after re-chromatography. The last lane contains the protein marker (M).

After re-chromatography a ceramic hydroxyapatite (CHT) chromatography step was performed. Upon binding of GBD to the hydroxyl apatite resin, the column material was subjected to an increasing phosphate gradient ranging from 0 to 500 mM of Sorensen's phosphate buffer of pH 6.5 over an elution volume of 55 mL (Figure 4-26). The corresponding SDS-Gel of the peak fractions shows protein of sufficient purity for the setup of crystallization trials (Figure 4-27).

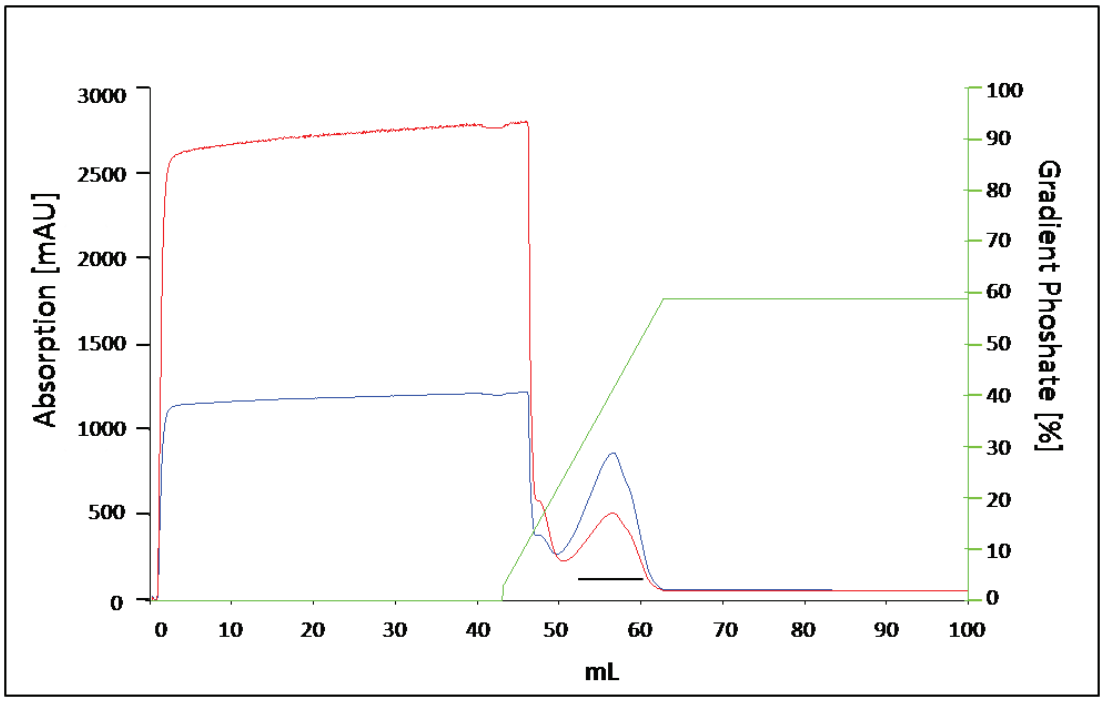

Figure 4-26: Ceramic hydroxyapatite (CHT) chromatography. The elution fractions of the preceding Ni-NTA affinity chromatography step were buffered at pH 6.5 and further purified via a ceramic hydroxyapatite column (HR 5/5). The chromatogram shows the absorption at two different wavelengths at λ = 280 nm (blue line) and λ = 260 (red line). The green line indicates the phosphate gradient. The black bar marks the pooled elution fractions.

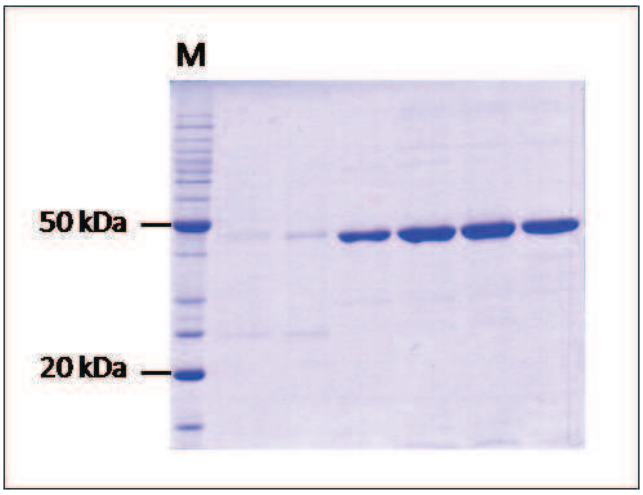

Figure 4-27: SDS-PAGE of CHT-chromatography peak fractions. First lane: protein standard (M), second lane: elution fractions.

Attempts to further purify the protein by GPC in order to remove possible aggregates resulted in the loss of the protein on the particular columns, most likely due to strong

interactions with the column matrix material, which often consists of cross-linked polysaccharides like dextran.

4.B.4 Crystallization of the GBD

The fractions containing purified GBD were pooled and concentrated (ultrafiltration; molecular weight cut-off of 10 kDa) to approximately 5 mg/mL. Adding of 10 % of glycerol turned out to be necessary in order to reach a protein concentration of about 5 mg/mL, otherwise the protein precipitated almost entirely.

Initial crystallization trials with the GBD were set up in 96-well sitting drop trays. The drops consisted of 100 nL protein (5 mg/mL, in 80 mM NaCl and 10 mM Tris, pH 7.8) mixed with 100 nL reservoir solution and were equilibrated against 70 nL reservoir buffer at 20°C.

Crystal formation of the GBD has been observed under two conditions of approximately 2000 crystallization conditions tested. These two conditions consisted of:

1) 2.0 M $AmSO_4$ and 0.1 M Na-acetate pH 4.6
2) 0.2 M $AmSO_4$, 0.1 M Mes pH 6.5 and 30 % PEG 5000 MME

The first condition produced elongated roundish crystals of small size up to 20 µm in the longest dimension. Subsequent attempts to reproduce this crystal form failed. The work on this crystal form was thus discontinued.

Small initial crystals grown from the second condition were reproduced and increased in size by microseeding to a maximum size of about 275 x 150 x 150 µm (Figure 4-28). These crystals were used in X-ray experiments (section 4.B.5).

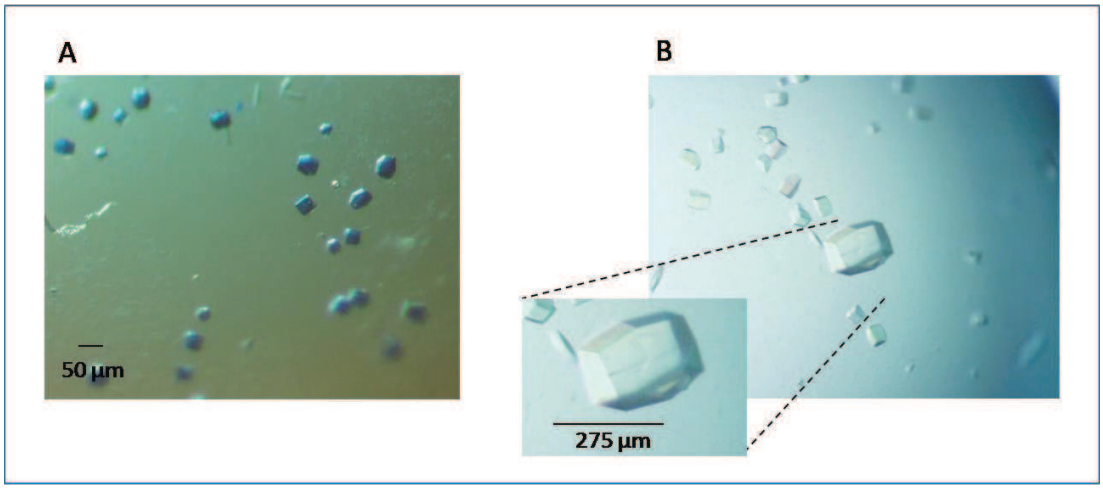

Figure 4-28: Crystals of the GBD. Grown from 0.2 M $AmSO_4$, 0.1 M Mes pH 6.5 and 30 % PEG 5000 MME. Crystals are colored with Izit Crystal dye (Hampton), indicating protein crystals (A). Crystals grown after microseeding (B).

4.B.5 X-ray data collection and space group determination of GBD crystals

As the crystals of GBD had been grown in 30 % PEG 5000 MME, the crystals were directly flash cooled at 100 K for data collection without any additional cryoprotectant. An X-ray diffraction data set of 360 images was collected at beamline ID29 at the ESRF (Grenoble, France). The oscillation angle was 0.5°, with an exposure time of 1 s per image. However, the best crystal gave only highly anisotropic diffraction patterns with resolution of 3.6 – 3.9 Å in one direction and 5.8 – 6.2 Å in the other direction (Figure 4-29).

Corresponding data collection statistics are summarized in Table 4-6. Analysis of the reciprocal lattice revealed no systematic absence along any axis. Indexing of GBD diffraction data was achieved using the program packages HKL2000 and XDS (Otwinowski and Minor, 1997; Kabsch, 2010), indicating a tetragonal symmetry. Preliminary X-ray analysis of these tetragonal crystals revealed unit cell dimensions of a = b = 86.00 Å and a rather long c-axis of 651.29 Å.

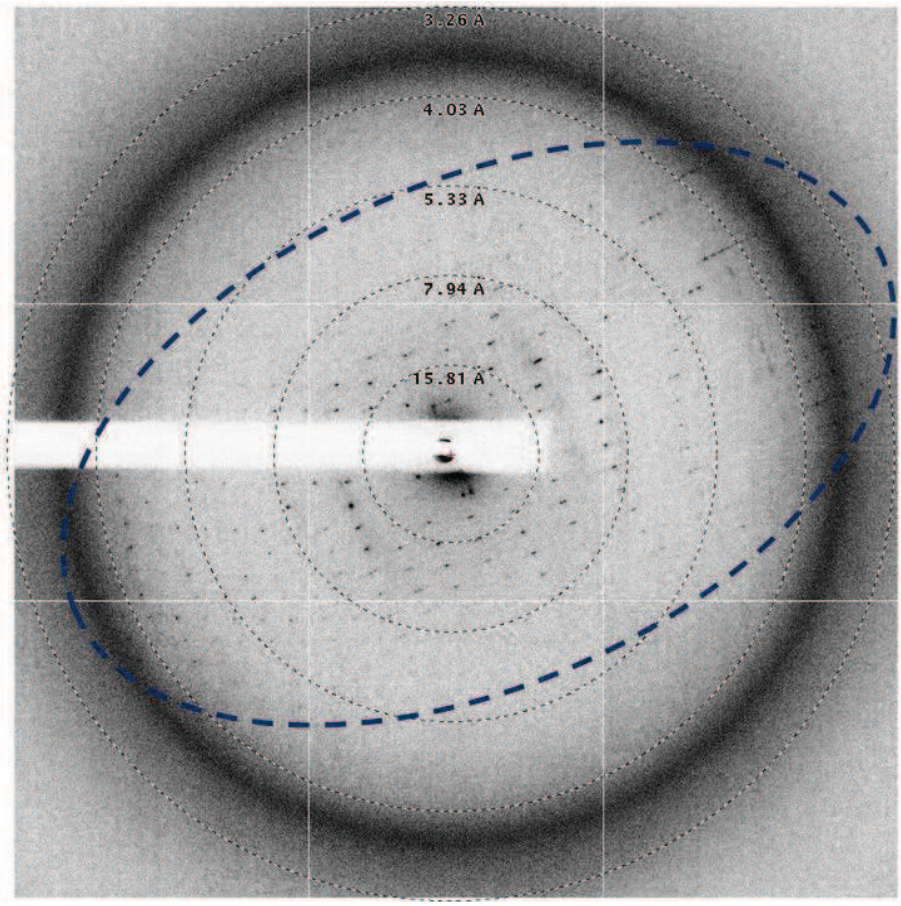

Figure 4-29: Diffraction image of a GBD crystal showing the anisotropic diffraction pattern.

Table 4-6: Summary of crystallographic data collection

	GBD
Wavelength (Å)	0.95002
Spacegroup	P422
Unit-cell parameters (Å,°)	a = b = 86.0, c = 651.29
	α = β = γ = 90.0
Resolution range (Å)	45.0 - 3.9
	45.0 - 4.7
No. of unique reflections	21607
Redundancy	6.9 (7.2)
Completness (%)	99.9 (97.4)
(I/σ(I))	3.1 (1.5)
R_{merge} (%)	23.0 (59.6)

Values in parentheses are for the highest resolution shell.

Matthews probability analysis (Matthews, 1968; Kantardjieff and Rupp, 2003) based on the molecular weight of 46.30 kDa for the GBD indicates there to be three to seven molecules per asymmetric unit (Table 4-7) for packing in space group P422.

Table 4-7: Calculated solutions of Matthews probabilities for GBD.

Number of molecules per asymmetric unit	Matthews Coefficient [Å^3/Da]	% solvent
3	4.27	71.18
4	3.20	61.57
5	2.56	51.97
6	2.13	42.36
7	1.83	32.76

4.B.6 Molecular Replacement

Diffraction data were integrated with the XDS package (Kabsch, 2010). Scaling was done first with SCALA (ccp4 Suite, 1994) but resulted in very low diffraction limits of about 4.7 Å due to the anisotropy of the data. Thus, anisotropic scaling was performed using the NIH Diffraction Anisotropy Server (http://www.doe-mbi.ucla.edu/~sawaya/anisoscale/) (Strong *et al*., 2006) with resulting resolution cutoffs between 3.7 and 4.6 Å.

To solve the structure by molecular replacement, several structures from the PDB were used as starting model. Three models were highlighted by the PHYRE server as having structural homology with the GBD. The PDB codes were 2bib (phosphorylcholine esterase from *Streptococcus pneumoniae*), 2g7c (*Clostridium difficile* Toxin A fragment), and 2wwc (choline binding domain of *Streptococcus pneumonia* autolysin C-Lyta). Each of the above mentioned individual search models (partially truncated) were used in attempts to solve the structure of the GBD by the use of MR. The structure of the choline binding domain of *S. pneumonia* autolysin C-Lyta (Figure 4-30) i.e. is formed by a series of repeats of so called choline-binding modules (CBM), which share a high homology to the YG repeats of glucansucrases (overall identity 32 %). Regarding the structural characteristics of these choline-binding modules, analysis of C-Lyta reveals a composition of repeating loop-β-hairpin structures.

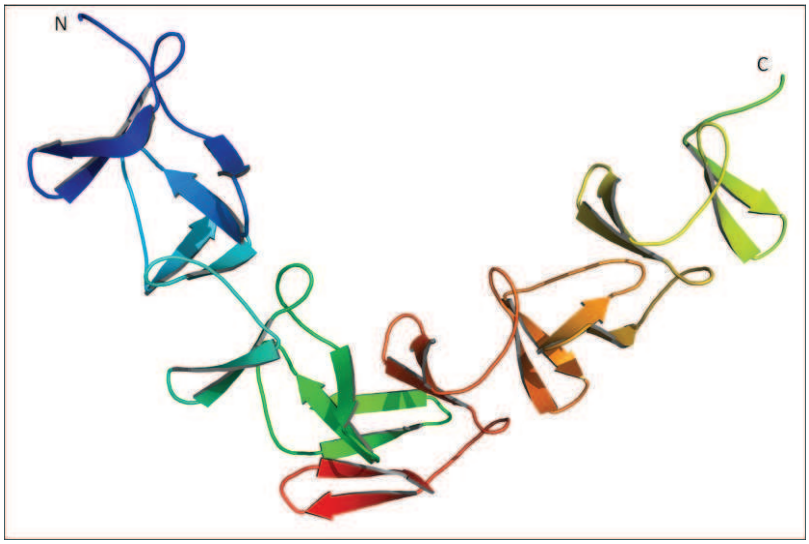

Figure 4-30: Choline binding domain of the major autolysin (C-Lyta) from *Streptococcus pneumonia*. The structure comprises a CBM of 11 repeat units.

Among each other, single repeats share a homology of up to 56 % (Figure 4-31). These single repeats were also used as starting models in molecular replacement attempts.

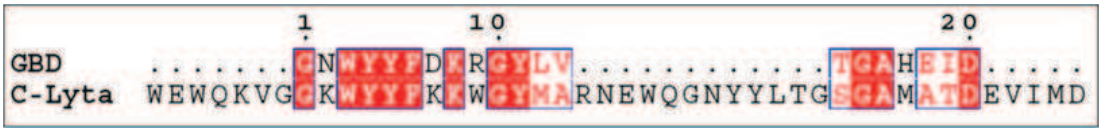

Figure 4-31: Alignment of the first YG repeat of GBD from *S. oralis* with a choline-binding module of C-Lyta from *S. pneumonia*.

Different search algorithms were applied (Patterson function superposition, maximum likelihood) using programs such as MOLREP (Vagin and Teplyakov, 1997), PHASER (McCoy *et al*, 2005) and AMoRe (Navaza, 1994), but none of the above models provided a suitable solution. Due to the anisotropic and low resolution data, structure solution of this data set was finally abandoned. Generation of high quality diffracting crystals will be necessary.

5 DISCUSSION

5.A Levansucrase SacB from *Bacillus megaterium*

The studies on levansucrase SacB from *B. megaterium* were performed in cooperation with Arne Homann and Sven Götze from the group of Prof. Dr. Jürgen Seibel (University of Würzburg, Institute of Organic Chemistry) and with Dr. Rebekka Biedendieck and Martin Gamer from the group of Prof. Dr. Dieter Jahn (TU Braunschweig, Department of Microbiology). The protein SacB with a relative molecular mass of 52 kDa was identified by Dr. Rebekka Biedendieck (2006). Together with Martin Gamer, she generated all mutants of *sacB* from *B. megaterium*, while Arne Homann and Sven Götze analyzed the product spectra of the wild-type and variant SacB. Since these results are an integral part of this thesis, they are mentioned and discussed in the following.

The crystal structure of levansucrase SacB from *B. subtilis (*Meng and Fütterer, 2003), the first structure of a GH 68 family member, revealed many details about this family of enzymes. In particular the complex structures with sucrose and raffinose gave insights about the amino acids located in the active site and involved in hydrolysis of sucrose. However, detailed knowledge about the transfructosylation process and residues involved in this process regarding oligo- and polysaccharide synthesis is still lacking for fructosyltransferases. In this thesis the crystal structures of five variants of SacB from *B. megaterium* have been determined at resolutions between 1.75 and 2.0 Å, in order to get more insights into levansucrase transfructosylation mechanism. The three-dimensional structures of SacB variants Y247A, Y247W, N252A, D257A and K373A display the same five-bladed β-propeller architecture as its homologue *B. subtilis* levansucrase enzyme. A funnel-like opening provides access to the deep negatively charged pocket towards the molecular surface. By superposition of SacB with its homologue from *B. subtilis*, the active site, containing the catalytic triad and further residues directly involved in the binding of sucrose have been identified at the bottom of the central cavity of the

β-propeller. Residues Asp95, Glu352 and Asp257 are proposed to compose the catalytic triad of levansucrase SacB from *B. megaterium* (Figure 4-9), with Glu352 acting as the acid/base catalyst and Asp95 as the nucleophile (Homann *et al*., 2007). Site-directed mutagenesis of these residues has been performed. As expected, it was shown that the substitution of Asp95 and Asp257 with alanine inhibited the enzyme activity (Homann, 2009). Residues involved in the sucrose binding constitute the -1 (fructosyl residue) and +1 (glucosyl residue) sugar binding subsites (nomenclature according to Davis *et al*., 1997). The cleavage of the substrate sucrose takes place between subsites -1 and +1. The enzyme forms a covalent intermediate with the fructose moiety of the cleaved sucrose at subsite -1. Subsequently, the fructose moiety is coupled to the acceptor molecule, whose fructose components are also bound to the enzyme at a number of subsites (-2, -3, -4…), depending on the length of the acceptor molecule (Homann *et al*, 2007).

5.A.1 SacB variants N252A and R370A point towards surface elements influencing polysaccharide synthesis

Homann and co-workers (2007) identified Asn252 as a residue playing an important role in the transfructosylation process of the levansucrase SacB from *B. megaterium*. The substitution of Asn252 with alanine or glycine abrogates the polysaccharide production at all, while the K_m and k_{cat} values were not affected and maintain in the wild-type range (K_m 4.1 mM, k_{cat} 1480 s^{-1}). Thin layer chromatography (TLC) analysis of the product spectrum showed that catalysis of these variants was switched from mainly polysaccharide synthesis to short-chain oligosaccharide synthesis and hydrolysis (Homann, 2009). As a result, SacB variant N252A formed oligosaccharides of up to three fructosyl units (nystose). Interestingly, the crystal structure of variant N252A shows that the substituted Asn252 is not located in the active site of the enzyme (Figure 4-10). The biochemical data along with the structural data show that N252A is located in subsite +2 (Figure 4-9), analogous to Asn242 of the homologous levansucrase SacB from *B. subtilis*, whose crystal structure was solved in complex with sucrose (PDB code: 1pt2) and raffinose (PDB code: 3byn) (Meng and Fütterer, 2003; 2008).

The crucial question arises which effect is responsible for the abrogation of polymer formation in variant N252A. Is Asn252 directly involved in interactions with the growing

fructan chain or does a mutation of Asn252 have an impact on the active site conformation and indirect effects are responsible for the abrogation of the polysaccharide production? The structure of SacB variant N252A shows an intact active site compared to D257A and K373A (Figure 4-10). None of the residues of the catalytic triad changed its conformation, indicating that indirect effects can be excluded. Supporting the crystallographic data, the kinetic parameters of variant N252A are not significantly different compared to wild-type SacB. The hydrolytic activity of N252A remains in the wild-type range and the variant is still able to cleave sucrose. These data feed the assumption that Asn252 is directly involved in interactions with the growing fructan chain. Due to its position in the +2 subsite it is suggested that the residue stabilizes the third fructosyl unit of the growing fructosyl chain and directs it as an acceptor substrate in the optimal position for further transfructosylation steps. The substitution of Asn252 to N252A switches the enzymatic reaction from poly- to oligosaccharide formation conserving a high enzyme activity by changing from a processive to a non-processive mechanism. As a result, the enzyme is not able anymore to synthesize oligosaccharides with more than three sugar units.

Kinetic studies and the characterization of the product spectrum also indicated the crucial role of Arg370 for the fructan polymerization in SacB from *B. megaterium* (Homann *et al.*, 2007). Arg370 is located in the +1 subsite and as well as Asn252. It is found that Arg370 is strictly conserved in fructosyltransferases from Gram positive bacteria such as *B. megaterium*. Structural comparison of the levansucrases SacB from *B. megaterium* and LsdA from *G. diazotrophicus* (Martinez-Fleites *et al*., 2005) enables insights into the functional role of Asn252 and Arg370. Superposition of the crystal structure of fructosyltransferase LsdA with SacB reveals no structural differences for the catalytic triad, but for some other regions in its vicinity. LsdA contains a histidine (His419) at the equivalent position instead of an arginine (Arg370, *B. megaterium*) (Figure 5-1). This His419 is found to be conserved in fructosyltransferases from Gram negative bacteria (Homann, 2009). Furthermore, the position of the strictly conserved Asn252 (*B. megaterium*) shows a high variability in Gram negative bacteria, demonstrated by the superposition of the corresponding region which shows that Asn252 is not superimposable with Asn306 (Figure 5-1). These structural differences in subsite +1 and +2 probably explain the differences in the product spectrum. Fructosyltransferases from

Gram negative bacteria synthesize mainly the β-(2,1)-linked trisaccharide 1-kestose, whereas fructosyltransferases from Gram positive bacteria mainly produce polysaccharides. This idea is supported by the observation that the product spectrum of N252A reveals no more polysaccharide production because Asn252 stabilizes the third fructosyl unit of the growing fructosyl chain. As a result, the polymerization process is abrogated and the product consists merely of three sugar units.

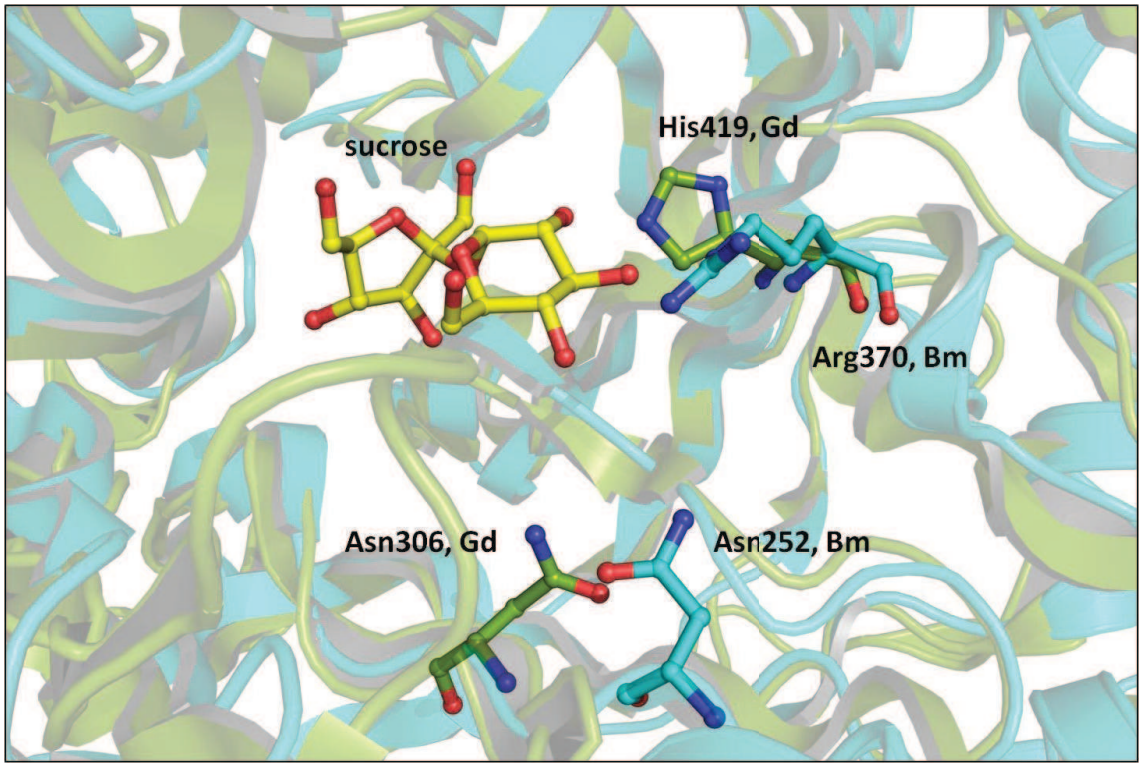

Figure 5-1: Local alignment and identification of residues Asn252 and Arg370 from the levansucrase SacB from *B. megaterium* crucial for polymer synthesis. SacB from *B. megaterium* (Bm, cyan) and LsdA from *G. diazotrophicus* (Gd, green) (Martinez-Fleites, 200) with modeled sucrose (yellow) in the active site are shown.

These results raise the question whether there are some other surface motifs influencing the polysaccharide synthesis of SacB? On the basis of structure alignments with the levansucrases SacB from *B. subtilis* and the related LsdA from *G. diazotrophicus*, eight additional amino acids with potential impact on the polysaccharide synthesis mechanism of SacB were identified. The amino acid functionality was knocked out by an exchange to alanine or partially maintained by the exchange to functionally similar amino acids. In order to elucidate the conformation and substrate affinity of the active site of the SacB variants, the kinetic parameters were determined and the product spectra were analyzed

(see sections 5.A.2 and 5.A.3) (Homann, 2009). Based on the results of this analysis, variants K373A, Y247A and Y247W were selected for crystallization and structure-function analyses.

5.A.2 SacB K373A abrogates the polysaccharide synthesis between subsite +4 and +5

Comparative kinetic analysis on wild-type SacB with variants K373A and K373R indicated that residue Lys373 has an important impact on the transfructosylation process of SacB. The exchange of Lys373 to alanine and arginine resulted in different product formation. HPAEC analysis of the product spectrum from K373A shows, that the formation of tri- and tetrasaccharides is lowered while the synthesis of penta- and hexasaccharides is slightly enhanced. However, the polysaccharide formation is completely stopped. The exchange to alanine leads to shorter oligosaccharides than the exchange to arginine. In contrast, variant K373R synthesized oligosaccharides of up to 7 units (Homann, 2009). Interactions between the arginine and the amino acid network as well as carbohydrate units still seem to be possible although not the whole range of oligosaccharides can be synthesized compared to the wild-type enzyme.

Consistent with the abrogated polysaccharide synthesis, K373A has an increased hydrolysis activity of almost 33 % (Homann, 2009). To clarify the role of Lys373 in the polysaccharide synthesis mechanism, the structure of SacB variant K373A was determined at a resolution of 1.75 Å. Lys373 is supposed to be located in subsite +4/+5. Superpositions with the structures of variants D257A and N252A show an intact active site. The functionally important amino acid residues are perfectly superimposable (Figure 4-10). Due to these structural data, again, indirect effects on the position of other residues in the active site can be excluded.

5.A.3 SacB Y247A abrogates the polysaccharide synthesis between subsites +8 and +9

The importance of Y247 to the transfructosylation process of SacB from *B. megaterium* was determined by the exchange to tryptophan, isoleucine and alanine, respectively.

The exchange of tyrosine in position 247 to alanine led to the formation of short oligosaccharides. HPAEC analysis of these oligosaccharides synthesized showed that the amount of octa- and nonasaccharides was slightly enhanced and the transfructosylation was totally abrogated after 9 fructosyl units. In contrast, the exchange of tyrosine to tryptophan resulted in the same oligosaccharide pattern as the wild-type SacB. Consistent with the kinetic parameters (K_m 2.1 mM, k_{cat} 2653 s^{-1}), variant Y247W totally maintained the wild-type activity and substrate affinity. The hydrolysis activity of variant Y247A was enhanced by 10 % while the hydrolysis products of Y247W were only slightly increased compared to wild-type SacB (5 %). To answer the question whether the conformation of the active site is influenced by the exchange of tyrosine in position 247, the X-ray structures of the SacB variants Y247A and Y247W were solved at a resolution of 2.0 and 1.9 Å, respectively. Superposition with variant D257A shows a conserved active site architecture of the catalytic triad as well as the functional important amino acids Arg370 and Asn252 (Figure 4-11). The intact active site was already implicated by a maintained K_m and k_{cat} of Y247A which is similar to the wild-type SacB.

The reason for the observation that Y247W forms the same oligosaccharide pattern as the wild-type may be that an unpolar favored π-π-stacking mechanism is possible with tyrosine as well as tryptophan in this location but not with alanine or isoleucine. The superposition of Y247W with D257A (Figure 4-11) confirms the key supporting role of Tyr247 in the oligo- and polysaccharide-forming “assembly line” (Figure 5-2).

5.A.4 Polysaccharide synthesis is controlled by distinct surface motifs

The structural data of variants N252A, K373A, Y247A, Y247W and D257A point towards a possible surface arrangement for the binding of an acceptor fructosyl chain. Residues Asn252, Lys373 and Tyr247 form a platform for a possible stabilization of the acceptor fructan chain. Clear subsites can be assigned to every exchanged amino acid (Figure 5-2). The biochemical data along with the structural data show that Asn252 is located in subsite +2, whereas Lys373 (subsite +4/+5) and Tyr247 (+8/+9) are even further outside the active site.

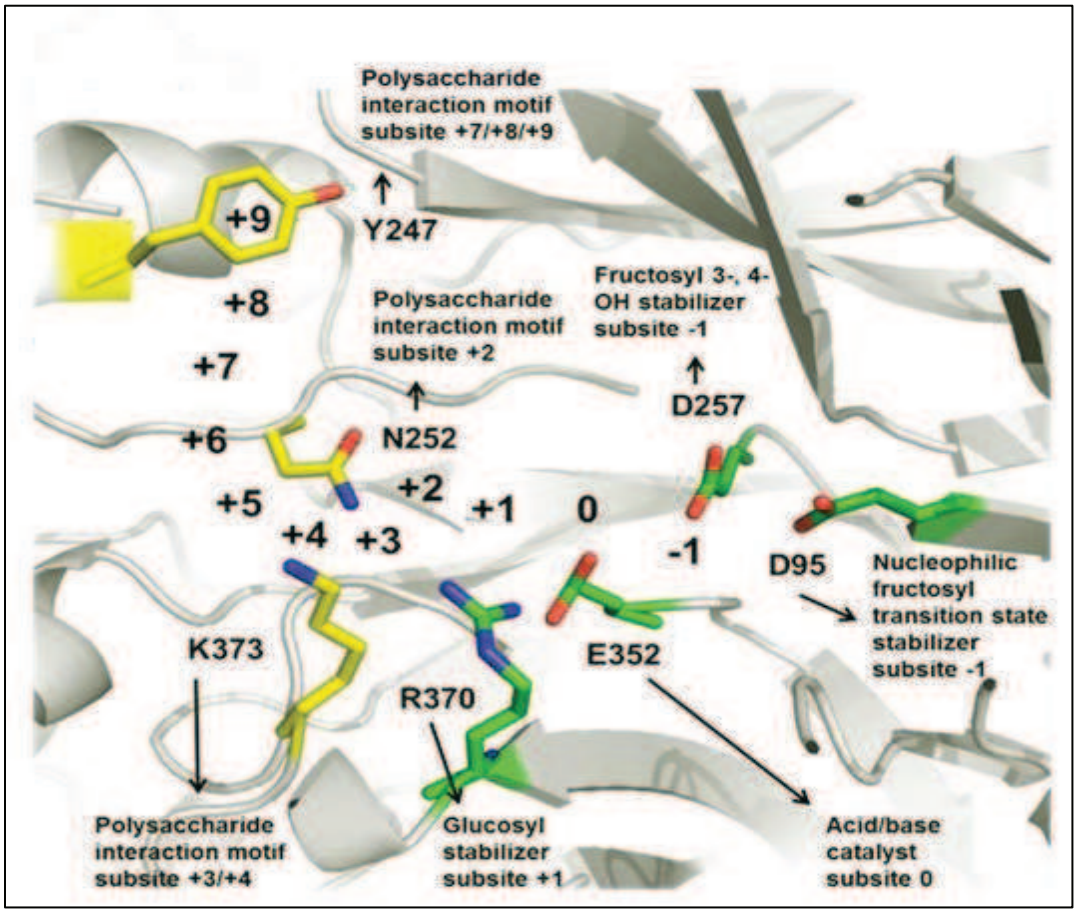

Figure 5-2: The surface localization of the polysaccharide interaction motifs of the levansucrase SacB from *Bacillus megaterium*. The amino acids interactions with the polysaccharide chain indicated are assigned according to the HPAEC analyses performed by Homann (2009).

All exchanged amino acids are located on the surface of SacB, indicating the interaction with the growing oligosaccharide chain. HPAEC analyses of the variant`s oligofructoside synthesis patterns, performed by Arne Homann (Homann *et al*., 2009) show definite abrogations of the polymerization process depending on the location of the mutated amino acid. Variants K373A, N252A and Y247A synthesize unique mixtures of oligosaccharides of different clearly distinguishable chain lengths correlating to their location on the surface of SacB (Figure 5-2). However, it remains unclear if the oligosaccharide chain has a conformation that places more than one fructosyl unit in one subsite of the enzyme or if the oligosaccharide chain interacts with even more amino acids on the surface of the fructosyltransferase. Most likely, there is a co-operation between the enzyme`s surface interaction with and the interaction between the fructosyl units of the oligosaccharide chain.

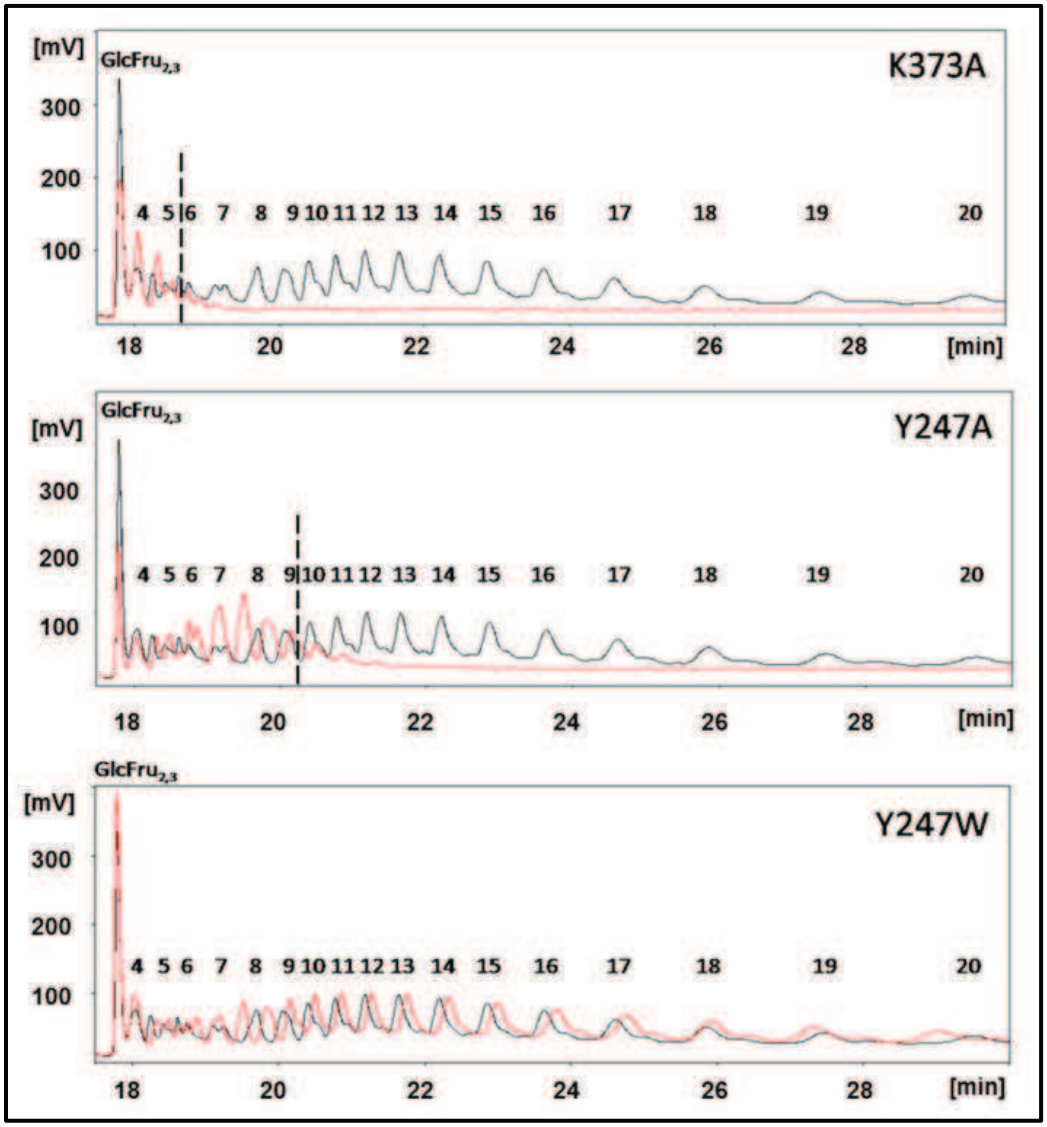

Figure 5-3: HPAEC analysis of the oligosaccharides synthesized by wildtype enzyme (black) and the respective variant (red). The peaks are assigned according to previous carbohydrate standard measurements (Homann *et al*., 2009).

5.A.5 Identification of the PEG molecules and probable explanation for failure to obtain structures of complex

Structural data of complexes between SacB and its fructosyl acceptors would have been very helpful for the investigation and understanding of the enzyme´s transfructosylation process. In order to generate these protein-ligand complexes, several co-crystallization-

and soaking experiments with different potential ligands have been applied. Despite numerous data sets that have been collected, processed, and evaluated, for no SacB variant a crystal structure could have been obtained with an unambiguously defined substrate, inhibitor or fructosyl acceptor, respectively. However, an elongated electron density was observed in the active site pocket of the crystallized enzymes. The crystallization solution used for SacB variants contained the precipitating agent PEG 1000. Attempts were made to fit PEG molecules into the density. PEG molecules with different molecular masses between 100 and 300 kDa matched the density well. It is well known that commercially available PEG reagents contain always a mixture of PEG molecules of different molecular masses, which would explain the electron density for a lower PEG mass than used in the crystallization solution. Another possibility is that only parts of the PEG molecules are well-ordered in the crystal and thus not the whole molecule is visible in the electron density map.

The presence of the bound PEG molecules was neither expected nor observed or reported in earlier structural studies on fructosyltransferases. The possibility that the PEG molecules mimic the fructosyl acceptor in the active site is unlikely. The observed hydrophobic interactions between the enzyme and the PEG molecules (Figure 4-15) show only interactions with residues which are not proposed to be involved in the transfructosylation process. However, the presence of the PEG molecules in the active site might avoid the accommodation of the added soaking substances. Additionally, the K_m value of 6.6 mM measured for the wild-type enzyme (Homann *et al*., 2007) indicates in general a very low substrate affinity compared with other enzymes *e.g*. SacB from *B. subtilis* with 13.5 to 40 mM according to BRENDA database (Schomburg *et al*., 2002). This could also explain the difficulties to remove the bound PEG fragments from the binding site of the crystallized protein. All attempts to crystallize SacB without a PEG containing solution failed (see sections 4.A.9, 4.A.10 and 4.A.11). On the contrary, Meng and Fütterer were able to obtain ligand bound enzyme crystals of SacB homologue from *B. subtilis* which possesses a sequence identity of 73 %. It is noteworthy that Fütterer did not choose the standard hanging or sitting drop method for the crystallization of SacB from *B. subtilis* (Meng and Fütterer, 2003). In contrast to his previous crystallographic work, published in numerous publications (Ahn *et al*., 2001; Wilson *et al*., 2003; Dover *et al*., 2004), Fütterer perfomed crystallization of SacB from *B. subtilis* by several steps of

dialysis without the use of PEG or any other precipitating agent. This observation supports the assumption that a PEG-free enzyme is necessary to obtain ligand-bound crystal structures of SacB from *B. subtilis* as well as from *B. megaterium*.

5.B Glycosyltransferase R from *Streptococcus oralis*

The interest in understanding the molecular mechanisms of enzymes associated with oligo- and polysaccharide production has pushed the efforts in solving structures of enzymes of the glucansucrase type. However, despite the enormous economical and pharmaceutical potential of these enzymes, no three-dimensional structure is available for a member of GH 70 family so far. The determination of the first one and so far only three-dimensional structure of a GH 70 family glucansucrase was reported in 2006 by Prof. Bauke Dijkstra (University of Groningen) on several conferences, but the corresponding data are still neither published nor available.

Glycosyltransferase R (GtfR) from *S. oralis* is a dextran-producing glucansucrase (E.C. 2.4.1.5), belonging to glycoside hydrolase family 70 (Hellmuth *et al*., 2008). Primary structure analyses revealed that most of glucansucrases possess a C-terminal domain, characterized by specific repeats called YG repeats. Several studies demonstrated that the C-terminal domain of these enzymes is involved in glucan binding. The exact role of the GBD in glucansucrase catalysis remains unresolved.

In this study the GBD of *S. oralis* could successfully be over-produced in *E. coli* and purified to a final yield of about 7 mg per L of culture. The choice of the domain boundaries for a structurally uncharacterized protein for crystallization purposes is always difficult. However, the selected domain boundaries for the GBD derived from comparison of sequence alignments (Figure 4-23) and the analysis of YG repeats led to the production of soluble protein, indicating at least the maintenance of structurally indispensable regions.

5.B.1 A probable explanation for the failed structure determination

Attempts to solve the structure of the GBD by the method of MR failed so far. In general, successful use of MR primarily depends on two factors: 1) the quality of the starting model of the protein available and 2) the nature of the crystal symmetry, packing and diffraction quality. Taylor (2003) showed that a sequence identity of at least 25 % between the model structure and the target protein is desirable. However, the more important point is the structural similarity between the two structures for successful MR

calculations. In the case of the GBD, starting models had a sequence identity of about 32 %, indicating a challenging case for structure determination by MR. There is a rough inverse correlation between the r.m.s. deviation of atomic positions and the percentage sequence identity. Good models have low r.m.s. deviation from the target structure and generally, low r.m.s. deviation between two structures is indicated by high sequence identity. The reason for the fact that no suitable solution was found for the GBD is most likely based on the low homology between the search models and the GBD. The anisotropy and the low resolution limits of the data even complicated the attempts to solve the structure. Another approach involving experimental phasing methods which provide phase information independent of any starting model seems to be necessary and should be pursued in future studies.

5.B.2 A probable explanation for the anisotropic diffraction

X-ray diffraction analysis of optimized GBD crystals resulted in highly anisotropic data with resolutions between 3.6 – 3.9 Å and in one direction and 5.8 – 6.2 Å in the other direction (Figure 4-29). These results indicate two problems for the structure determination of the GBD. The first difficulty is the low resolution limit of about 3.6 Å even in the “best” direction. Nevertheless, in principle, the determination of a low resolution structure at this resolution is possible. However, the second difficulty concerns the anisotropy of the data, which strongly impairs the possible determination of the low resolution structure. Diffraction anisotropy arises when the number of lattice contacts is less in one cell direction than in another. Indexing of GBD diffraction data indicated unit cell dimensions of a = b = 86.0 Å and an extremely long c-axis of 651.3 Å. In this case, the unit cells would have to be in a proper order in the c-direction over tenfold the range compared to the a/b direction in order to reach the same resolution.

Another explanation for the anisotropy in the diffraction pattern could simply be due to the shape of the crystals. The intensity of diffraction is a function of the volume of diffracting matter that is hit by the X-ray beam. Therefore, diffraction anisotropy is often accompanied by platy or needle-shaped crystals. However, the GBD crystals grew to a size of about 275 x 150 x 150 µm, indicating that the crystal shape of GBD plays only a negligible role for the explanation of its anisotropic behavior.

Once the structure of the GBD is solved it will be possible, “a posteriori”, to give a plausible explanation of the anisotropy.

Future strategies should focus on the generation of high quality crystals with different crystal forms, which may show a different crystal packing and more isotropic diffraction properties.

REFERENCES

Ahn, S., Milner, A. J., Fütterer, K., Konopka, M., Ilias, M., Young, T. W. and White, S. A. (2001). The "open" and "closed" structures of the type-c inorganic pyrophosphatases from *Bacillus subtilis* and *Streptococcus gordonii*. *J. Mol. Biol*. 313: 797-811.

Alberto, F., Bignon, C., Sulzenbacher, G., Henrissat, B. and Czjzek, M. (2004). The three-dimensional structure of invertase from *Thermotoga maritime* reveals a bimodular arrangement and an evolutionary relationship between retaining and inverting glycosidases. *J. Biol. Chem*. 279: 18903-10.

Arguello-Morales, M. A., Remaud-Simeon, M., Pizzut, S., Sarcabal, P. and Monsan, P. (2000). Sequence analysis of the gene encoding alternansucrase, a sucrose glucosyltransferase from *Leuconostoc mesenteroides* NRRL B-1355. *FEMS Microbiol. Lett.* 182: 81-85.

Ausubel, F. M., Brent, R., Kingston, R. E., Moor, D. D., Seidman, J. G., Smith, J. A. and Struhl, K. (2007). Current protocols in molecular biology. John Wiley & Sons Inc.

Banas, J. A. and Vickermann, M. M. (2003). Glucan-binding proteins of the oral streptococci. *Crit. Rev. Oral Biol. Med*. 14: 89-99.

Beine, R., Moraru, R., Nimtz, M., Na`amnieh, S., Pawlowski, A., Buchholz, K. and Seibel, J. (2008). Synthesis of novel fructooligosaccharides by substrate and enzyme engineering. *J. Biotechnol.* 38: 33-41.

Berg, J. M., Tymoczko, J. L. and Stryer, L. (2007). *Biochemistry.*

Bergfors, T. (2003). Seeds to crystals. *J. Struct. Biol.* 142: 66-76.

Biedendieck, R. (2006). TU Braunschweig, PhD thesis. *Bacillus megaterium*: Versatile Tools for Production, Secretion and Purification of Recombinant Proteins

Broadbent, J. R., McMahon, D. J., Welker, D. L., Oberg, C. J. and Moineau, S. (2003). Biochemistry, genetics, and applications of exopolysaccharide production in *streptococcus thermophilus*: a review. *J. Dairy Sci*. 86: 407-23.

Bryson, K., McGuffin, L. J., Marsden, R. L., Ward, J. J., Sodhi, J. S. and Jones, D. T. (2005). Protein structure prediction servers at university college London. *Nucleic Acids Res*. 33: W36-8.

Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V. and Henrissat, B. (2009). The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Res.* 37: 233-8.

Cerning, J. (1990). Exocellular polysaccharides produced by lactic acid bacteria. *FEMS*

Microbiol. Rev. 87: 113-30.

Cerning, J., Bouillanne, C., Desmazeaud, M. J. and Landon, M. (1986). Isolation and characterization of exocellular polysaccharide produced by *Lactobacillus bulgaricus*. *Biotechnol. Lett*. 8: 625-28.

Chambert, R. and Gonzy-Treboul, G. (1976). Levansucrase of *Bacillus subtilis*. Characterization of a stabilized fructosyl-enzyme complex and identification of an aspartyl residue as the binding site of the fructosyl group. *Eur. J. Biochem*. 71: 493-508.

Chambert, R., Gonzy-Treboul, G. and Dedonder, R.(1974). Kinetic studies of levansucrase of *Bacillus subtilis*. *Eur. J. Biochem*. 41:285-300.

Chambert, R. and Petit-Glatron, M. F. (1991). Polymerase and hydrolase activities of *Bacillus subtilis* levansucrase can be separately modulated by site-directed mutagenesis. *Biochem. J*. 279: 35-41.

Chen, V. B., Arendall, W. B. 3rd., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. and Richardson, D. C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* 66: 12-21.

Chuankhayan, P., Hsieh, C. Y., Huang, Y. C., Hsieh, Y. Y., Guan, H. H., Hsieh, Y. C., Tien, Y. C., Chen, C. D., Chiang, C. M. and Chen, C. J. (2010). Crystal structure of *Aspergillus japonicus* fructosyltransferase complex with donor/acceptor substrates reveal complete subsites in the active site for catalysis. *J. Biol. Chem.* 285: 23251-64.

Cole, C., Barber, J. D. and Barton, G. J. (2008). The Jpred secondary structure prediction server. *Nucleic Acids Res*. 36: W197-201.

Coligan, J. E. (2003). Short protocols in protein science. A compendium of methods from Current protocols in protein science. John Wiley & Sons Inc.

Crescenzi, V. (1995). Microbial polysaccharides of applied interest: ongoing research activities in europe. *Biotechnol. Prog*. 11: 251-9.

Crittenden, R. G. and Playne, M. J. (1996). Purification of food-grade oligosaccharides using immobilized cells of *Zymomonas mobilis*. *Appl. Microbiol. Biotechnol*. 58: 297-302.

Daguer, J. P., Geissmann, T., Petit-Glatron, M. F. and Chambert, R. (2004). Autogenous modulation of the *Bacillus subtilis* sacB-levB-yveA levansucrase operon by the levB transcript. *Microbiology.* 150: 3669-79.

Davis, G. J., Wilson, K. S. and Henrissat, B. (1997). Nomenclature for sugar-binding subsites in glycosyl hydrolases. *Biochem. J.* 321: 557-59.

De Vuyst, L. and Degeest, B. (1999). Heteropolysaccharides from lactic bacteria. *FEMS Microbiol. Rev*. 23: 153-77.

DeLano, W. L. (2002). The PyMOL Molecular Graphics System. DeLano Scientific, San Carlos, CA, USA.

Dover, L. G., Corsino, P. E., Daniels, I. R., Cocklin, S. L., Tatituri, V., Besra, G. S. and Fütterer, K. (2004). Crystal structure of the TetR/CamR family repressor *Mycobacterium tuberculosis* EthR implicated in ethionamide resistance. *J. Mol. Biol.* 340: 1095-1105.

Edman, P. and Begg, G. (1967). A protein sequenator. *Eur. J. Biochem.* 1: 80-91.

Emsley, P., Lohkamp, B., Scott, W. G. and Cowtan, K. (2010). Features and development of Coot. *Acta Cryst. D. Biol. Crystallogr.* 66: 486-501.

Endo-Streeter, S. T. (2009). Duke University, PhD thesis. Structural Studies of *Arabidopsis thaliana* Inositol Polyphosphate Multi-Kinase.

Feingold, D. S., Avigad, G. and Hestrin, S. (1956). The mechanism of polysaccharide production from sucrose. 4. Isolation and probable structures of oligosaccharides formed from sucrose by levansucrase system. *Biochem. J.* 64: 351-61.

French, A. D. and Waterhouse, A. L. **(**1993). Chemical structure and characteristics, p. 41–82. In Suzuki, M. and Chatterton, N. J. (ed.), Science and technology of fructans. CRC Press Inc., Boca Raton, Fla.

Fujiwara, T., Hoshino, T., Ooshima, T., Sobue, S. and Hamanda, S. (2000). Purification, characterization, and molecular analysis of the gene encoding glucosyltransferase from *Streptococcus oralis*. *Infect. Immun*. 68: 2475-83.

Geresh, S., Mamontov, A. and Weinstein, J. (2002). Sulfation of extracellular polysaccharides of red microalgae: preparation, characterization and properties. *J. Biochem. Biophys. Meth*. 50: 179-187.

Gibson, G. R. and Roberfroid, M. B. (1995). Dietary modulation of the human colonic microbiota: introducing the concept of prebiotics. *J. Nutr.* 125: 1401-12.

Giffard, P. M. and Jacques, N. A. (1994). Definition of a fundamental repeating unit in streptococcal glucosyltransferase glucan-binding regions and related sequences. *J. Dent. Res.* 73: 1133-41.

Gill, S. C. and von Hippel, P. H. (1982). Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem.* 182: 319-26.

Gouet, P., Robert, X. and Courcelle E. (2003). ESPript/ENDscript: Extracting and rendering sequence and 3D information from atomic structures of proteins. *Nucleic Acids Res.*
31: 3320-3.

Grobben, G. J., Van Casteren, W. H., Schols, H. A., Oosterveld, A., Sala, G., Smith, M. R., Sikkema, J. and De Bont, J. A. M. (1997). Analysis of the exopolysaccharides produced by *Lactobacillus delbrueckii* subsp. bulgaricus NCFB 2772 grown in

continuous culture on glucose and fructose. *Appl. Microbiol. Biotechnol*. 48:516-20.

Hamada, S. and Slade, H. D. (1980). Biology, immunology, and cariogenicity of *Streptococcus mutans*. *Microbiol. Rev*. 44: 331-84.

Heger, A. and Holm, L. (2000). Rapid automatic detection and alignment of repeats in protein sequences. *Proteins*. 41: 224-37.

Heinig, M. and Frishman, D. (2004). STRIDE: a web server for secondary structure assignment from known atomic coordinates of proteins. *Nucleic Acids Res*. 32: W500-2.

Hellmuth, H., Hilringhaus, L., Höbbel, S., Kralj, S., Dijkhuizen, L. and Seibel, J. (2007). Highly efficient chemoenzymatic synthesis of novel branched thiooligosaccharides by substrate direction with glucansucrases. *Chembiochem*. 8: 273-6.

Hellmuth, H. Wittrock, S., Kralj, S., Dijkhuizen, L., Hofer, B. and Seibel, J. (2008). Engineering the glucansucrase GTFR enzyme reaction and glycosidic bond specificity: toward tailor-made polymer and oligosaccharide products. *Biochemistry*. 47: 6678-84.

Hempel, D. (2006). Integration gen- und verfahrenstechnischer Methoden zur Entwicklung biotechnologischer Prozesse Sonderforschungsbereich 578 – Vom Gen zum Produkt. *Chemie Ingenieur Technik* 78: 187-92.

Heng, J., Randall, A. Z., Sweredoski, M. J. and Baldi, P. (2005). SCRATCH: a protein structure and structural feature prediction server. *Nucleic Acids Res*. 33: W72-6.

Henrissat, B. (1991). A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem. J.* 280: 309-16.

Henrissat, B. and Davies, G. (1997). Structural and sequence-based classification of glycoside hydrolases. *Curr. Opin. Struct*. *Biol*. 7: 637-44.

Hernandez, L., Arrieta, J., Menendez, C., Vazquez, R., Coego, A., Suarez, V., Selman, G., Petit-Glatron, M. F. and Chambert, R. (1995). Isolation and enzymic properties of levansucrase secreted by *Acetobacter diazotrophicus* SRT4, a bacterium associated with sugar cane. *Biochem. J.* 309: 113-18.

Hestrin, D., Feingold, D. S. and Avigad, G. (1956). The mechanism of polysaccharide production from sucrose. 3. Donor-acceptor specificity of levansucrase from *Aerobacter levanicum*. *Biochem. J*. 64: 340-51.

Hestrin, S. and G. Avigad. (1958). The mechanism of polysaccharide production from sucrose. 5. Transfer of fructose to C-1 of aldose by levansucrase. *Biochem. J*. 69: 388-98.

Homann, A., Biedendieck, R., Götze, S., Jahn, D. and Seibel, J. (2007). Insights into polymer versus oligosaccharide synthesis: mutagenesis and mechanistic studies of a novel levansucrase from *Bacillus megaterium*. *Biochem. J*. 407: 189-98.

Homann, A. (2009). TU Braunschweig, PhD thesis. En Route to Tailor-made

Oligosaccharides – Chemo-enzymatic Synthesis and Physiological Functions of Novel Carbohydrate Structures.

Hosono, A., Lee, J., Ametani, A., Natsume, M., Hirayama, M., Adachi, T. and Kaminogawa, S. (1997). Characterization of a water-soluble polysaccharide fraction with immunopotentiating activity from *Bifidobacterium adolescentis*. *Biosci. Biotechnol. Biochem*. 61: 312-16.

Johnson, K. F. (1999). Synthesis of oligosaccharides by bacterial enzymes. *Glycoconj. J.* 2: 141-46.

Jolly, L., and Stingele, F. (2001). Molecular organization and functionality of exopolysaccharides gene clusters in lactic acid bacteria. *Int. Dairy J*. 11: 733-45.

Jolly, L., Newell, J., Porcelli, I., Vincent, S. J. and Stingele, F. (2002). *Lactobacillus helveticus* glycosyltransferases: from genes to carbohydrate synthesis. *Glycobiology*. 12: 319-27.

Kabsch, W. (2010). Integration, scaling, space-group assignment and post refinement. *Acta Cryst. D.* 66: 133-44.

Kabsch, W. and Sander, C. (1983). Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*. 22: 2577-637.

Kang, K. S. and Cottrell, I. W. (1979). Polysaccharides, p. 417-81. In H. J. Pepler and D. Perlman (ed), Microbial technology: microbial processes, 2nd ed., Academic Press, Inc, New York.

Kantardijeff, K. A. and Rupp, B. (2003). Matthews coefficient probabilities: improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals. *Protein Sci*. 12: 1865-71.

Kato, C. and Kuramitsu, H. K. (1991). Molecular-basis for the association of glucosyltransferases with the cell-surface or oral streptococci. *FEMS Microbiol. Lett*. 79: 153-58.

Kelley, L. A. and Sternberg, M. J. (2009). Protein structure prediction on the web: a case study using the phyre server. *Nat. Protoc*. 4: 363-71.

Kenne, L., and Lindberg, B. (1983). Bacterial polysaccharides in polysaccharides. New York: Academic Press Vol. 2.

Kim, N. Y., Kim, H. G., Kim, Y. H., Chung, I. S. and Yang, J. M. (2008). Expression and characterization of human N-acetylglucosaminyltransferases and alpha 2,3-sialyltransferase in insect cells for in vitro glycosylation of recombinant erythropoietin. *J. Microbiol. Biotechnol.* 18: 383-91.

Kingston, K. B., Allen, D. M., and Jacques, N. A. (2002). Role of the C terminal YG repeats of the primer-dependent streptococcal glucosyltransferase, GtfJ, in binding to

dextran and mutan. *Microbiology* 148: 549–58.

Kleywegt, G. J., Henrick, K., Dodson, E. J. and van Aalten, D. M. (2003). Pound-wise but penny-foolish: How well do micromolecules fare in macromolecular refinement? *Structure* 11: 1051-59.

Kojic, M., Vujcic, M., Banina, A., Cocconcelli, P., Cerning, J. and Topisirovic, L. (1992). Analysis of exopolysaccharide production by *Lactobacillus casei* CG11, isolated from cheese. *Appl. Environ. Microbiol.* 58: 4086-88.

Koshland, D. E. and Stein, S. S. (1954). Correlation of bond breaking with enzyme specificity. Cleavage point of invertase. *J. Biol. Chem*. 208: 139-48.

Kralj, S. (2004). Rijksuniversiteit Groningen, PhD thesis. Glucansucrases of Lactobacilli: Characterization of genes, enzymes and products synthesized.

Kralj, S., van Geel-Schutten, G. H., Rahaoui, H., Leer, R. J., Faber, E. J., van der Maarel, M. J. and Dijkhuizen, L. (2002). Molecular characterization of a novel glucosyltransferase from *Lactobacillus reuteri* strain 121 synthesizing a unique, highly branched glucan with α-(1→4) and α-(1→6) glucosidic bonds. *Appl. Environ. Microbiol*. 68: 4283-91.

Kumar, A. S., Mody, K. and Jha, B. (2007). Bacterial polysaccharides – a perception. *J. Basic Microbiol.* 47(2): 103-17.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227: 680-85.

Lammens, W., Le Roy, K., Schroeven, L., Van Laere, A., Rabijns, A. and Van den Ende, W. (2008). Structural insights into glycoside hydrolase family 32 and 68 enzymes: functional implication. *J. Exp. Bot.* 60: 727-40.

Lamzin, V. S. and Wilson, K. S. (1993). Automated refinement of protein models. *Acta Cryst. D. Biol. Crystallogr.* 63: 759-68.

LeBrun, E. and Van Rapenbusch, R. (1980). The structure of *Bacillus subtilis* levansucrase at 3.8 Å resolution. *J. Biol. Chem.* 255: 12034-36.

Leslie, A. G. W. (1992). Recent changes to the MOSFLM package for processing film and image plate data. *Joint CCP4+ESF-EAMCB Newsletter on Protein Crystallography.*

Lis, M., Shiroza, T. and Kuramitsu, H. K. (1995). Role of the C-terminal direct repeating units of the *Streptococcus mutans* glucosyltransferase-S in glucan binding. *Appl. Env. Microbiol*. 61: 2040-42.

Loesche, W. J. (1986). Role of S*treptococcus mutans* in human dental decay. *Microbiol. Rev.* 50: 353-80.

Lovell, S. C., Davis, I. W., Arendall, W. B. 3rd, de Bakker, P. I., Prisant, M. G.

Richardson, J. S. and Richardson, D. C. (2003). Structure validation by C alpha geometry: phi, psi and C beta deviation. *Proteins*. 50: 437-50.

Macura, D. and P. M. Townsley. 1984. Scandiavian ropy milk - identification and characterization of endogenous ropy lactic streptococci and their extracellular secretion. *J. Dairy Sci*. 67: 735-44.

Martinez-Fleites, C., Ortiz-Lombardia, M., Pons, T., Tarbouriech, N., Taylor, E. J., Arrieta, J. G., Hernandez, L. and Davies, G. J. (2005). Crystal structure of levansucrase from the Gram-negative bacterium *Gluconacetobacter diazotrophicus*. *Biochem. J*. 390: 19-27.

Matthews, B. W. (1968). Solvent content of protein crystals. *J. Mol. Biol.* 33: 491-97.

McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C. and Read, R. J. (2005). Likelihood-enhanced fast translation functions. *Acta Crystallogr. D. Biol. Crystallogr*. 61: 458-64.

Medzhitov, R. and Janeway, C. A. Jr. (1997). Innate immunity: impact on the adaptive immune response. *Curr. Opin. Immunol*. 9: 4-9.

Meng, G. and Fütterer, K. (2003). Structural framework of fructosyl transfer in *Bacillus subtilis* levansucrase. *Nat. Struct. Biol*. 10: 935-41.

Meng, G. and Fütterer, K. (2008). Donor substrate recognition in the raffinose-bound E342A mutant of fructosyltransferase *Bacillus subtilis* levansucrase. *BMC Struct. Biol*. 8: 16.

Messerschmidt, A. (2007). X-Ray crystallography of biomacromolecules: A practical guide. Wiley-VCH Verlag. 304 pp.

Meulenbeld, G. H. and Hartmans, S. (2000). Transglycosylation by *Streptococcus mutans* GS-5 glucosyltransferase-D: acceptor specificity and engineering of reaction conditions. *Biotechnol. Bioeng*. 70: 363-69.

Monchois, V., Reverte, A., Remaud-Simeon, H., Monsan, P. and Willemot, R. M. (1998). Effect of *Leuconostoc mesenteroides* NRRL B-512F dextransucrase carboxy-terminal deletions on dextran and oligosaccharide synthesis. *Appl. Environ. Microbiol*. 64: 1649.

Monchois, V., Willemot, R. M. and Monsan, P. (1999). Glucansucrases: mechanism of action and structure-function relationships. *FEMS Microbiol. Rev*. 23: 131–51.

Moulis, C., Joucla, G., Harrison, D., Fabre, E., Potocki-Verones, G., Monsan, P. and Remaud-Simeon, M. (2006). Understanding the polymerization mechanism of glycoside-hydrolase family 70 glucansucrases. *J. Biol. Chem*. 281: 31254-267.

Murshudov, G. N., Vagin, A. A. and Dodson, E. J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D. Biol. Crystallogr.* 53: 240-55.

Navaza, J. (1994). AMoRe: an automated package for molecular replacement. *Acta Cryst. A* 50: 157-63.

Niesen, F. H., Berglund, H., & Vedadi, M. (2007). The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat. Prot., 2*(9), 2212-2221.

Oliveira, M. A., Guimaraes, B. G., Cussiol, J. R., Medrano, F. J., Gozzo, F. C. and Netto, L. E. (2006). Structural insights into enzyme-substrate interaction and characterization of enzymatic intermediates of organic hydroperoxide resistance protein from *Xylella fastidiosa*. *J. Mol. Biol.* 359: 433-45.

Ortiz-Soto, M. E., Rivera, M., Rudino-Pinera, E., Olvera, C. and Lopez-Munguia, A. (2008). Selected mutations in *Bacillus subtilis* levansucrase semi-conserved regions affecting its biochemical properties. *Protein Eng. Des*. *Sel.* 21: 589–95.

Otwinowski, Z. and Minor, W. (1997). Macromolecular Crystallography Part A: [20] Processing of X-ray diffraction data collected in oscillation mode. *Meth. Enzymol*. 276: 307-26.

Ozimek, L. K., Kralj, S., van der Maarel, M. J. and Dijkhuizen, L. (2006). The levansucrase and inulosucrase enzymes of *Lactobacillus reuteri* 121 catalyse processive and non-processive transglycosylation reactions. *Microbiology*. 152: 1187-96.

Parker, R. B. and Creamer, H. R. (1971). Contribution of plaque polysaccharides to growth of cariogenic microorganisms. *Arch. Oral Biol*. 16: 855-62.

Parolis, L. A., Parolis, H., Kenne, L., Meldal, M. and Bock, K. (1998). The extracellular polysaccharide of *Pichia* (Hansenula) *holstii* NRRL Y-2448: the phosphorylated site chains. *Carbohydr. Res.* 309: 319-25.

Prasad, J. C., Comeau, S. R., Vajda, S. and Camacho, C. J. (2003). Consensus alignment for reliable framework prediction in homology modeling. *Bioinformatics*. 19: 1682-91.

Quiocho, F. A. (1986) Carbohydrate-binding proteins: tertiary structures and protein-sugar interactions 2. *Annu. Rev. Biochem.* 55: 287-315.

Roberfroid, M. B. (2000). Prebiotics and probiotics: are they functional food? *Am. J. Clin. Nutr.* 71: 1660S-4S.

Roberfroid, M. B., Van Loo, J. A. E. and Gibson, G. R. (1998). The bifidogenic nature of chicory inulin and its hydrolysis products. *J. Nutr.* 128: 11-9.

Roberts, C. M., Fett, W. F., Osman, S. F., Wijey, C., O'Connor, J. V. and Hoover D. G. (1995). Exopolysaccharide production by *Bifidobacterium longum* BB-79. *J. Appl. Bacteriol*. 78:463-468.

Rupp, B. (2009). Biomolecular Crystallography: Principles, Practice, and Application to structural biology. Garland Science 1st edition.

Rye, C. S. and Withers, S. G. (2000). Glycosidase mechanisms. *Curr. Opin. Chem. Biol*. 4: 573-80.

Sambrook, J. and Russell, D. W. (2000). Molecular Cloning - a laboratory manual. Cold Spring Harbor Laboratory Press.

Schlegel, H., Ed. (1992). Allgemeine Mikrobiologie, Thieme.

Schomburg, I., Chang, A. and Schomburg, D. (2002). BRENDA, enzyme data and metabolic information. *Nucleic Acids Res*. 30: 47-9.

Schreuder, H. A., Groendijk, H., van der Laan, J. M. and Wierenga, R. K. (1988). The transfer of protein crystals from their original mother liquor to a solution with a completely different precipitant. *J. Appl. Cryst. Biol. Crystallogr* 21: 426-29.

Schwarz, A., Brecker, L. and Nidetzky, B. (2007). Acid-base catalysis in *Leuconostoc mesenteroides* sucrose phosphorylase probed by site-directed mutagenesis and detailed kinetic comparison of wild-type and Glu237 → Gln mutant enzymes. *Biochem. J.* 403: 441-9.

Seibel, J., Hellmuth, H., Hofer, B., Kicinska, A. M. and Schmalbruch, B. (2006). Identification of new acceptor specifities of glycosyltransferase R with the aid of substrate microarrays. *Chembiochem*. 7: 310-20.

Seo, E. S., Lee, J. H., Park, J. Y., Han, H. J. and Robyt, J. F. (2005). Enzymatic synthesis and anti-coagulant effect of salicin analogs by using the *Leuconostoc mesenteroides* glucansucrase acceptor reaction. *J. Biotechnol*. 117: 31-8.

Sinclair, A. M. and Elliott, S. (2005). Glycoengineering: the effect of glycosylation on the properties of therapeutic proteins. *J. Pharm. Sci*. 94: 1626-35.

Song, D. D. and N. A. Jacques. 1999. Purification and enzymic properties of the fructosyltransferase of *Streptococcus salivarius* ATCC 25975. Biochem. J. 341: 285-291.

Stingele, F., Neeser, J. R. and Mollet, B. (1996). Identification and characterization of the eps (Exopolysaccharide) gene cluster from *Streptococcus thermophilus* Sfi6. *J. Bacteriol*. 178: 1680-90.

Strong, M., Sawaya, M. R., Wang, S, Philips, M., Cascio, D. and Eisenberg, D. (2006). Towards the structural genomics of complexes: crystal structure of a PE/PPE protein complex from *Mycobacterium tuberculosis*. *Proc. Natl. Acad. Sci*. USA, 103: 8060-65.

Stura, E. A. and Wilson, I. A. (2003). Applications of the streak seeding technique in protein crystallization. *J. Cryst. Growth.* 110: 270-82.

Sutherland, I. W. (1972). Bacterial exopolysaccharides. *Adv. Microbiol. Physiol*. 8: 143-212.

Sutherland, I. W. (1990). Biotechnology of microbial exopolysaccharides. Cambridge,

UK: Cambridge University Press.

Sutherland, I. W. (1999). Microbial polysaccharide products. *Biotechn. Genet. Eng. Rev.* 16: 217-29.

Suwannarangsee, S., Moulis, C., Potocki-Veronese, G., Monsan, P., Remaud-Simeon, M. and Chulalaksananukul, W. (2007). Search for a dextransucrase minimal motif involved in dextran binding. *FEBS Lett.* 581: 4675-80.

Taylor, G. (2003). The phase problem. *Acta Crystallogr. D. Biol. Crystallogr.* 59: 1881-90.

Tieking, M., Ehrmann, M. A., Vogel, R. F. and Gänzel, M. G. (2005) Molecular and functional characterization of a levansucrase from the sourdough isolate *Lactobacillus sanfranciscensis* TMW 1.392. *Appl. Microbiol. Biotechnol*. 66: 655-63.

Tsuchiya, H. M., Hellman, N. N. and Koepsell, H. J. (1953). Factors a¡ecting molecular weight of enzymatically synthesized dextran. *J. Am. Chem. Soc.* 75: 757-58.

Ungar, D. (2009). Golgi linked protein glycosylation and associated diseases. *Semin. Cell. Dev. Biol.*

Vagin, A. and Teplyakov, A. (2010). Molecular Replacement with MOLREP. *Acta Crystallogr. D Biol. Crystallogr.* 66: 22-5.

Vaguine, A. A., Richelle, J. and Wodak, S. J. (1999). SFCHECK: a unified set of procedures for evaluating the quality of macromolecular structure-factor data and their agreement with the atomic model. *Acta Crystallogr. D. Biol. Crystallogr.* 55: 191-205.

Van Casteren, W. H., de Waard, P., Dijkema, C., Schols, H. A., and Voragen A. G. (2000). Structural characterisation and enzymic modification of the exopolysaccharide produced by *Lactococcus lactis subsp*. cremoris B891. *Carbohydr. Res*. 327: 411-22.

van Geel-Schutten, G. H., Faber, E. J., Smit, E., Bonting, K., Smith, M. R., Ten Brink, B., Kamerling, J. P., Vliegenthart, J. F. and Dijkhuizen, L. (1999). Biochemical and structural characterization of the glucan and fructan exopolysaccharides synthesized by the *Lactobacillus reuteri* wild-type strain and by mutant strains. *Appl. Environ. Microbiol*. 65: 3008-14.

van Hijum, S. A., Kraji, S., Ozimek, L. K., Dijkhuizen, L. and van Geel-Schutten, I. G. (2006). Structure-function relationships of glucansucrase and fructansucrase enzymes from lactic acid bacteria. *Microbiol. Mol. Biol. Rev.* 70: 411-33.

Van Hijum, S. A. (2004). Rijksuniversiteit Groningen, PhD thesis. Fructosyltransferases of *Lactobacillus reuteri*: characterization of genes, enzymes, and fructan polymers

Varki, A., Cummings, R., Esko, J., Hudson, F., Gerald, H., and Jamex, M. (1999). Essentials of glycobiology 2nd edition, Cold Spring Habor Laboratory Press

Varki, A. (2006). Nothing in glycobiology makes sense except in the light of evolution. *Cell* 126: 841-5.

Velázquez-Hernández, M. L., Baizabal-Aguirre, V. M., Bravo-Patino, A. Cajero-Juarez, M., Chavez-Moctezuma, M. P. and Valdez-Alarcon, J. J. (2009). Microbial fructosyltransferases and the role of fructans. *J. Appl. Microbiol*. 106: 1763-78.

Vickermann, M. M., Sulavik, M. C., Minick, P. E. and Clewell, D. B. (1996). Changes in the carboxy-terminal repeat region a¡ect extracellular activity and glucan products of *Streptococcus gordonii* glucosyltransferase. *Infect. Immun*. 64: 5117-28.

Vujičić-Žagar, A. (2007). Rijksuniversiteit Groningen, PhD thesis. Structural and functional investigations of *Lactobacillus reuteri* glucansucrase: ... with crystallographic studies on an α-amylase and a prolyl endoprotease from *Aspergillus niger*.

Vriend, G. (1990). WHAT IF: a molecular modeling and drug design program. *J. Mol. Graph.* 8: 52-6.

Walker, G. J. (1978). Dextrans. *Int. Rev. Biochem*. 16: 75-126.

Wallace, A. C., Laskowski, R. A. and Thornton, J. M. (1995). LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Eng*. 8: 127-34.

Whitfield, C., and W. J. Keenleyside. (1995). Regulation of expression of group IA capsular polysaccharides in Escherichia coli and related extracellular polysaccharides in other bacteria. *J. Ind. Microbiol*. 15: 361-71.

Wilson, R. A., Maughan, W. N., Kremer, L., Besra, G. S. and Fütterer, K. (2003). The structure of *Mycobacterium tuberculosis* MPT51 (FpbC1) defines a new family of non-catalytic α/β hydrolases. *J. Mol. Biol.* 335: 519-530.

Yun, J. W. (1996). Fructooligosaccharides – occurrence, preparation and application. *Enzyme Microb. Technol.* 19: 107-17.

Zhu, D. Y., Zhu, Y. Q., Xiang, Y. and Wang, D. C. (2005). Optimizing protein crystal growth through dynamic seeding. *Acta Crystallogr. D Biol. Crystallogr.* 61: 772-5.

Zwart, P. H., Grosse-Kunstleve, R. W. and Adams, P. D. (2005). Xtriage and Fest: automatic assessment of X-ray data and substructure structure factor estimation. ccp4 newsletter 43.

Zwart, P. H., Grosse-Kunstleve, R. W., Lebedev, A. A., Murshudov, N. and Adams, P. D. (2008). Surprises and pitfalls due to (broken) symmetry. ccp4 newsletter 46.

FIGURES

TABLES

DANKSAGUNG

Prof. Dr. Dirk Heinz danke ich herzlich für die Zeit in seiner Arbeitsgruppe und die hervorragenden Arbeitsbedingungen, die er gewährleistet hat. Sein Interesse an meinem Thema, sowie seine Diskussionsbereitschaft weiß ich sehr zu schätzen.

Dr. Gunhild Layer danke ich sehr herzlich für die freundliche Übernahme des Zweitgutachtens.

Bei Prof. Dr. Dieter Jahn bedanke ich mich herzlich für die freundliche Übernahme des Vorsitzes der Prüfungskommission, sowie für die Kooperation im Rahmen des SFB.

Bei Prof. Dr. Jürgen Seibel, Prof. Dr. Petra Dersch, Arne Homann, Sven Götze, Rebekka Biedendieck, Dr. Bernd Hofer, Jens Schneider, Malte Timm und Martin Gamer möchte ich mich für die hervorragende Kooperation und die Diskussionsbereitschaft bedanken.

Dr. Björn Klink, Dr. Joachim Reichelt und Prof. Dr. Wolf-Dieter Schubert gilt ein besonderer Dank für ihre unendliche Geduld und Unterstützung im Graphikraum und in den zahllosen Seminaren.

Der Arbeitsgruppe von Prof. Ralf Ficner möchte ich für die Möglichkeit zur Nutzung des Röntgengerätes und ihre bereitwillige Unterstützung danken. Auch allen Beamline Scientists am DESY in Hamburg, am BESSY in Berlin und am ESRF in Grenoble möchte ich für ihre Hilfe bei den zahlreichen Röntgenexperimenten danken.

Desweiteren bedanke ich mich bei Dr. Manfred Nimtz und seinen Teamkollegen Undine Felgenträger und Anja Meier für die Massenspektren samt sehr gründlicher Diskussion.

Bei Rita Getzlaff möchte ich mich für die zahlreichen N-terminalen Sequenzierungen bedanken.

Katrin Rand, Maike Rochon, Boris Grujic, Agnes Zimmer, Christin Holland, Nils Kuklik, Lilia Polle, Jens de Groot, Matthias Haffke, Ute Widow und natürlich Maike Bublitz möchte ich ganz besonders herzlich für die vielen kleinen und großen Unterstützungen während der Doktorarbeit und für die wundervolle Arbeitsatmosphäre danken.

Ein großes Dankeschön geht an meine Diplomandin Maria Ebbes. Die Zusammenarbeit hat mir viel Spaß gemacht und war auch für mich sehr lehrreich. Zudem danke ich meinen fleißigen Praktikantinnen Heike Laschin und Corinna Probst, die ebenfalls zum Gelingen der Arbeit beigetragen haben. Auch Heikes Katze und Kaninchen, die mich mit dem nötigen seeding-Material versorgt haben, danke ich sehr.

Weiterer Dank gilt Christine Bentz, Carina Büttner, Yvonne Carius, Felix Deluweit, John Dorian, Alexander Eberth, Davide Ferraris, Daniela Gebauer, Ute Grumer, Gregor Hagelüken, Claudia Hanko, Thomas Heidler, Joop van den Heuvel, Birgit Hofmann, Alexander Iphöfer, Nadine Konisch, Jörn Krauße, Björn Niebel, Stefanie Loss, Hartmut Niemann, Anja Menzel, Steffen Meyer, Edukondalu Mullapudi, Nick Quade, Carolin Schaper, Sabine Schmidt, Jörg Schulze, Stephanie Schulz, Barney Stinson, Kevin Walkling, Uwe Wengler, Ulrich Wiesand, Luisa Winkler, Thomas Wollert, Victor Wray, dem CLUB, sowie unseren Nachbarn, der CPRO Gruppe.

Meiner Mutter und meiner Schwester möchte ich für den bedingungslosen Rückhalt danken, ohne den diese Arbeit gar nicht möglich gewesen wäre.

Mein ganz besonderer Dank gilt meiner Freundin Susanne, die in allen Momenten für mich da war und mir die nötige Kraft und Motivation für den erfolgreichen Abschluß der Doktorarbeit gegeben hat.